Systems-Level Analysis of the Toll-like Receptor Network of Dendritic Cells

Dissertation Advisor: Dr. Nir Hacohen Nicolas Chevrier

Systems-Level Analysis of the Toll-like Receptor Network of Dendritic Cells

Abstract

Cells detect and respond to environmental changes using intracellular networks, and defects in the wiring of these networks contribute to diseases. For example, Toll-like receptors (TLRs) sense microbial molecules and trigger pathways critical for host defense. Genetic defects in components of the TLR and other pathogen-sensing pathways have been linked to human diseases. Hence, rational targeting of these pathways should help to manipulate immune responses associated with infections, autoimmunity, or vaccines. A fundamental challenge is to dissect the intracellular networks mobilized by pathogen-sensing pathways. Here we present approaches to dissect the TLR network of innate immune dendritic cells (DCs), focusing on two regulatory layers: signaling and transcription.

First, we present a strategy to systematically perturb candidate regulators and monitor cellular transcriptional responses. We apply this approach to derive regulatory networks that control the transcriptional response to TLR engagement by microbial molecules. Our approach revealed the regulatory functions of 125 transcription factors (TFs), chromatin modifiers, and RNA binding proteins, which enabled the construction of a network model consisting of 24 core regulators and 76 “fine-tuners” that help explain how TLR pathways achieve specificity.

Second, we report the systematic discovery of signaling components in TLR responses. By combining transcriptional profiling, genetic and small molecule perturbations, and phosphoproteomics, we uncover 35 signaling regulators, including 16 known members of the TLR signaling pathways. In particular, we find that Polo-like kinases (Plk) 2 and 4 are essential components of antiviral pathways *in vitro* and *in vivo* and activate a signaling branch involving a

dozen proteins, among which is Tnfaip2, a gene associated with autoimmune diseases but whose role was unknown.

Lastly, we expand these approaches to integrate functional and physical interactions linking the 'signaling-to-transcription' TLR network. By combining our perturbation-based approach with measurements of physical interactions, including phosphorylation, protein complexes, and TF binding to DNA, we uncover 30 signaling regulators mechanistically linked to 19 downstream TFs. The integration of these datasets into a model reveals the organization of the TLR response.

Overall, these studies illustrate the power of combining systematic measurements and perturbations to elucidate complex intracellular circuits and discover potential therapeutic targets.

Acknowledgements

I would like to express sincere thank to my advisor, Nir Hacohen. The continuous mentorship, support, opportunities, and engaging research environment Nir has provided have contributed greatly to helping my scientific and personal development.

I would like to thank all past and present members of the Hacohen lab. First, I can't give enough credit to Ido Amit, who was a great mentor during my first years in the lab. Other lab members to who I am greatly thankful for their continuous help along the way: Tom Eisenhaure, who has been a pillar of the lab since the beginning and made key contributions to this work; Raktima Raychowdhury, Chloe Villani, Arnon Arazi, Weibo Li, Liguo Wu, Alice Yuk, Will Pendergraft, and my fellow graduate students in the lab: Marciela DeGrace, Mark Lee, Matt Roy, Karolina Maciag, and Kelli Deering.

I am greatly indebted to Aviv Regev and her group. She has been instrumental in the realization of my scientific projects and a great support. I would especially like to thank Max Artyomov, Nir Yosef, and Manuel Garber, whose help and patience have been vital to the completion of this work.

Steven Carr and his group have had an invaluable impact on the success of this collaborative work. Steven has provided tremendous support and expertise. Our projects would simply not have been feasible without his group members Karl Clauser and Philipp Mertins. Philipp has been and remains a close collaborator without whom none of this work would have been possible.

I am also very grateful to the RNAi Consortium for their continuous support and help, and to the Broad Institute as a whole. The richness of scientific endeavors happening at this institute has greatly enhanced my learning and development as a scientist.

I would like to acknowledge each of my collaborators whose help made enormous contributions to my research: Alex Shalek (Hongkun Park lab), Matteo Iannacone and Elena Tonti (Ulrich von Andrian lab), Mark Ciaccio (Richard Jones lab).

Additionally, I am extremely thankful to my thesis committee members: Diane Mathis, Laurie Glimcher, Paul Anderson, and David Bartel for their continuing support and advice. I am also grateful to my thesis examiners: Jonathan Kagan, Mikael Pittet, and Jeroen Saeij.

I would also like to extend my sincere gratitude to the HMS immunology program for the fantastic opportunity of joining their Ph.D. program, especially Christophe Benoist for his initial support, and to the Boehringer Ingelheim Fonds for financial as well as scientific support.

I would like to address special thanks to all my former scientific mentors, chronologically: Pierre Offre, Shizuo Akira and Satoshi Uematsu, Bruce Beutler and Kasper Hoebe, Jean-Pierre Gorvel and Suzana Salcedo, and Eric Vivier and Sophie Ugolini. They have all been great sources of inspiration and their collective mentorship, training, and support allowed me to reach my Ph.D. training well prepared.

Je tiens à adresser mes sincères remerciements à mes parents, Michel and Caroline, mes soeurs, Johanna et Pauline, ainsi qu'à mes amis et mes proches.

Table of Contents

List of Figures

Chapter 1 – Systems-Level Analysis of Biological Processes

1.1. Overview

In this thesis work, we report three systems-level approaches to dissect pathogen-sensing pathways of innate immunity, focusing on two regulatory layers: signaling and transcription. Using these approaches in the context of dendritic cells (DCs), we discovered a host of Toll-like receptor (TLR) network components that participate in the control of both of these regulatory layers. Prior to detailing our findings, we provide in the following sections a brief overview of: (*i*) systems approaches to biological and immunological questions, (*ii*) the current knowledge of pathogen-sensing pathways, and (*iii*) the main goals and questions driving the work presented here.

1.2. Studying Complex Biological Systems

Cells detect and respond to environmental changes using intracellular networks. These networks arise from interactions among cellular components, including proteins, nucleic acids, and metabolites. Furthermore, the spatial and temporal arrangement of these components and the network that they comprise are key to appropriate cellular behavior. To reach a mechanistic understanding of these networks, a fundamental challenge is thus to identify and connect all cellular components, and to uncover the rules governing their design principles and functioning.

Conceptual Framework. The past decade has witnessed a fundamental transition from the study of individual components to modules emanating from many interacting components (Hartwell et al., 1999). This shift can be attributed to conceptual and technological advances

leading to what are now commonly denominated systems approaches (Ideker et al., 2001; Kitano, 2002). These approaches may vary in their implementation but largely rely on a similar conceptual framework: (*i*) identify all components of biological system; (*ii*) systematically perturb these components by deleting or modifying their properties or concentrations, and monitor the effect on the system's output; (*iii*) extract and connect the functional modules of the system to derive quantitative models (Ideker et al., 2001; Ross, 2008).

The sequencing of the human genome (Lander et al., 2001; Venter et al., 2001) has perhaps been the single-most important study in terms of driving these conceptual and technological changes. Following the completion of this project, much of the past decade has led to an exponential increase in studies using systems approaches across a variety of prokaryotic and eukaryotic models. Furthermore, recent advances in measurement tools now allow the capture of entire sets of cellular components.

The next challenge is now to dissect cellular circuitry (Lander, 2011). The technological advances enabling large-scale to complete listing of biological components (*e.g.*, transcripts, phosphorylated proteins, chromatin marks), have outpaced data analysis and integration that translate large-scale measurements into biologically meaningful information. To start tackling this challenge and reach a mechanistic understanding of biological networks and their constitutive modules, it is crucial to move toward integrative analyses across types of components (*e.g.*, RNA and proteins) and regulatory layers (*e.g.*, signaling and metabolic networks), and incorporating temporal and spatial variables (*e.g.*, short-term vs. evolutionary time scale).

Technological Framework. Many technological advances have propelled biology to the next level of understanding and helped starting to accomplish the vision delineated above. Over the

last twenty years, many technological developments and inventions have enabled the capture of entire sets of cellular components including nucleic acids, proteins, lipids, or metabolites. For example, complete genomic DNA sequences of microbes, plants, insects, or humans can be readily generated, if not already available. Beyond these now typical DNA sequencing projects, metagenomic measurements of the DNA content within communities of organisms such as bacteria are more and more widely applied. Next-generation sequencing approaches are also steadily replacing hybridization-based, microarray methodologies to monitor gene expression and discover novel transcripts or alternative splice forms. Finally, great advances in mass spectrometry instrumentation have been at the origin of “next-generation” proteomics, as opposed to previous methods that would rely, for example, on 2D electrophoresis of proteins based on mass and charge followed by identification of proteins contained within individual spots on the gel. Current mass spectrometry instruments allow the measurements of entire proteomes in simpler model organisms such as yeast, and with unprecedented depth of coverage in mammalian cells. Furthermore, enrichment of subproteomes such as phosphorylated, ubiquitinylated, or acetylated proteins using antibodies or chromatographic methods, has led to the investigation of many aspects of protein function, regulation, and modification within a cell that had remained out of reach for a long time.

In parallel to these advances in measurement technologies of cellular components, the toolbox for perturbing these components at nearly any level (*e.g.*, DNA, RNA, protein) has greatly expanded as well. For example, we have moved from gene knockouts and random mutagenesis as being the only feasible approaches to perturb or delete a gene, to conditional mutants (*e.g.*, in cell types of interest), to post-transcriptional silencing (*e.g.*, RNAi), or more recently to genome editing (e.g., zinc-finger nucleases). For human biology, the pendant for these perturbations may be seen in natural genetic variations across the human population, given that

human genomes exhibit on average ~20 inactivated genes per individual due to loss-of-function variants (MacArthur et al., 2012).

Overall, the implementation of this conceptual framework – in concert with the model and biological questions at hand – may help building the functional wiring of a cell. The complexity of mammalian systems together with remaining technical limitations still preclude us from connecting these functional wires, but recent advances in lower model organisms have demonstrated how this view may be attained to some extent. For example, a recent multi-lab effort focused on *Mycoplasma pneumoniae* – a bacterium with 689 proteins encoded in a ~800 kilobase genome – to perform detailed measurements of multiple regulatory layers including: protein-coding and non-coding RNA expression, RNA splicing, protein-protein interactions, and metabolic pathways (Ochman and Raghavan, 2009). This landmark analysis suggested that the level of complexity of the regulatory pathways in place in this "simpler" organism was higher than previously thought. Furthermore, it provides a data framework for many studies including systems-level analyses of this bacterium's physiology and evolution. Other examples come from studies using yeast, a unicellular eukaryotic organism. For example, genome-scale genetic interaction maps have proven useful in dissecting the functional redundancy underlying the functional wiring of cells and organisms (Costanzo et al., 2011). Systematic efforts for mapping these genetic interactions in *Saccharomyces cerevisiae*, using mutant libraries with double gene deletions, have led to enormous amounts of information regarding interactions between and within biological pathways, and thus, helped in starting to derive the functional map of a yeast cell (Costanzo et al., 2011). The few examples cited above demonstrate the power of systems-level approaches in elucidating the wiring of biological processes, and help laying the path to move forward with more complex mammalian systems. However, they also clearly remind us that the level of complexity of biological models, as simple as they may seem, is still high

enough that many more technological and conceptual advances will be required to generate truly functional maps of a cell.

1.3. Studying Complex Immunological Systems

Systems approaches have also made a steady foray into the study of the immune system (Benoist et al., 2006; Germain et al., 2011), including innate immune pathways (Zak and Aderem, 2009). Numerous reports have focused on various regulatory layers of immune cells, and here, we will focus on systematic studies of signaling and transcriptional layers, which are directly connected to the primary goals of this thesis work. Of note, other regulatory layers are equally important to the understanding of immune processes. For example, studies of the epigenome and associated regulatory pathways that are important for immune functions have also emerged in recent past (Northrup and Zhao, 2011). Human $CD4^+$ T cells are among the cell types that have been the most widely characterized at the epigenomic level (Northrup and Zhao, 2011). Indeed, several studies have collectively profiled the genome-wide distributions of many histone modifications and TF binding events, and are setting up the stage to many more similar studies. These studies have contributed to the basic understanding of chromatin biology and immune cell type division and differentiation.

Transcriptional Regulatory Layer. Studies of transcriptional responses and networks have perhaps led to most contributions of systems approaches to deciphering immune functions. This is likely due to the early advent of enabling technologies in transcriptional profiling, such as microarrays (Schena et al., 1995), compared to the more recent advances in protein-based profiling using mass spectrometry. It led to many contributions in fields ranging from pathogen sensing (Gilchrist et al., 2006; Jenner and Young, 2005; Litvak et al., 2009; Suzuki et al., 2009), to hematopoiesis (Novershtern et al., 2011), to human diseases such as autoimmunity

(Guiducci et al., 2010; Pascual et al., 2010) or tuberculosis (Berry et al., 2010), to vaccinology (Nakaya et al., 2011; Pulendran et al., 2010; Querec et al., 2009), and even to the Immunological Genome Project (Heng and Painter, 2008).

Host transcriptional responses to microorganisms and viruses have been extensively studied following the advent of microarrays (Jenner and Young, 2005). For example, some of the initial studies in this field led to the identification of IFN-induced genes (ISGs), which are regulated in host cells upon exposure to IFNs (Jenner and Young, 2005). These studies have thus highlighted common sets of components induced at the transcriptional level, either directly or indirectly, by pathogens. Depending on the host cell-type or pathogen considered, the sets of induced components may vary, but in all cases, it appeared clear that multiple transcriptional regulators were necessary to coordinate every aspect of these responses. Altogether, these studies paved the way to functional studies whereby host genes may be perturbed genetically (*e.g.*, knockout) or post-transcriptionally (*e.g.*, RNAi) and the result of such perturbation on pathogen-induced responses may be monitored by transcriptional profiling.

In the field of TLR biology, initial studies using TLR4 signaling induced by lipopolysaccharide (LPS; Gram-negative bacterial wall component) have generated important insights for the understanding of the transcriptional regulation occurring in pathogen sensing (Gilchrist et al., 2006; Litvak et al., 2009). For example, using microarray profiling to identify LPS-regulated genes in macrophages, these studies identified candidate transcriptional regulators by scanning for TF-binding sites in the promoter regions of LPS-regulated genes (Gilchrist et al., 2006; Litvak et al., 2009). The transcription factors ATF3 and C/EBPδ were identified as critical regulators downstream of the LPS sensor TLR4, and their roles were further investigated using functional (*i.e.*, knockouts) and physical (*i.e.*, chromatin immunoprecipitation) approaches. Overall, these studies highlighted the necessity for multiple TFs to work in concert as sub-circuits to finely

control LPS-regulated gene expression, raised the questions of how these TFs may be systematically identified and how their inter-relationships may be mapped.

Studies in other model systems have made important contributions to start tackling these challenges using integrated computational and experimental approaches. For example, a group combined transcriptional profiling together with functional studies using siRNA in a myeloid cell line, THP-1, in the context of stimulation with phorbol myristate acetate (PMA) (Suzuki et al., 2009). First, they scanned for TF binding motifs within close proximity of the promoter regions of PMA-regulated genes in their THP-1 cellular model system. Second, identified binding sites were used to infer the variation in activity of TFs over time upon PMA stimulation. Third, the 30 core TFs identified by these analyses were assembled in a predicted transcriptional regulatory network. Finally, the authors of this study were able to validate some of the predictions emanating from their network modeling by knocking down selected TFs from their core network using siRNAs. Altogether, this study in THP-1 cells and the work highlighted above in TLR4-mediated transcriptional regulation helped advancing our knowledge of transcriptional regulatory circuits. However, these studies remain limited by the use of TF binding sites which remains a poor predictor of TF-target gene relationships, and thus, limits *de novo* identification of TFs involved in a biological response to a stimulus since many TF's binding sites remain poorly defined or unknown.

As exemplified by the studies delineated above, transcriptional regulatory networks are critical in controlling acute cellular responses to a stimulus such as a pathogen or its molecular components. Transcriptional networks are also critical in driving differentiation processes such as hematopoiesis, and in maintaining differentiated cell types such as immune cell subsets. For example, the Immunological Genome Project has undertaken a multi-laboratory effort to generate genome-wide mRNA profiles of >200 immune cell types (Heng and Painter, 2008).

These datasets will undoubtedly be key to start dissecting the transcriptional networks operating the division of labor within cell types of the immune system. Another example of this view comes from a recent study that carried out genome-wide mRNA profiling of 38 human hematopoietic cell types (Novershtern et al., 2011). From these profiles, ~80 modules of co-expressed genes were identified across these 38 cell types. Furthermore, enriched *cis*-regulatory elements within each module highlighted TFs that are likely to contribute to the regulation of these modules. Using shRNA-mediated knockdowns, some of these TFs were validated as being functionally important in driving the differentiation of hematopoietic lineages *in vitro*. Altogether, this integrated computational and experimental study provided an initial draft for the complex transcriptional regulatory circuits driving the differentiation of human hematopoietic lineages. As stated above, because the binding sites for many TFs are still unknown, this study is likely to be underestimating the actual number of TFs necessary to finely control these differentiation programs.

Studies of transcriptional responses in immune cells have also helped with deciphering the mechanism of action and predicting the efficacy of therapies against infectious and non-infectious diseases. For example, recent studies have used transcriptional profiling in human blood immune cells to identify early gene regulatory events that correlate with effective vaccine regimens, as measured by antibody and T cell responses several weeks upon administration (Nakaya et al., 2011; Pulendran et al., 2010; Querec et al., 2009). Using a similar approach, others have highlighted a previously unrecognized role for type I interferons (IFNs) in the pathogenicity of tuberculosis, and furthermore, identified potential biomarkers for various stages of the disease (Berry et al., 2010). Another study, focused on systemic lupus erythematosus (SLE), suggested a mechanism that helps explain why many SLE patients show resistance to glucocorticoids, common treatment against autoimmune disorders such as SLE (Guiducci et al., 2010).

Collectively, the studies presented in this section exemplify how transcriptional regulation may be harnessed to generate advances in both basic understanding of immune processes, and in therapies aiming to manipulate these processes.

Signaling Regulatory Layer. In this section, we focus on the emergence of systems approaches applied to the understanding of signal transduction. The regulation of signaling processes is tightly connected to downstream transcriptional programs. Generally speaking, the propagation of signaling events is linked to changes in the post-translational modifications (PTMs) of proteins (*e.g.*, phosphorylation), in protein-protein interactions, and in protein concentrations (Choudhary and Mann, 2010). These various components of signaling regulatory layers can now be measured and studied in an unbiased fashion thanks to recent advances in mass spectrometry (MS)-based proteomics (Choudhary and Mann, 2010). These advances have spurred numerous findings in various immunological processes such as signaling regulation in dendritic cells (Luber et al., 2010), macrophages (Weintz et al., 2010), and T cells (Navarro et al., 2011), among other cell types (Fraser and Germain, 2009).

Large-scale measurements of PTMs using MS-based proteomics have perhaps led to one of the biggest data explosion in the field of signaling regulation. For example, analyses of phosphorylation levels at steady state and upon stimulation have identified tens of thousands of phosphosites across the eukaryotic proteome. In the field of TLR biology, the study of the TLR4-regulated phosphoproteome led to the identification of potentially new signaling modules as players of the TLR pathways (Weintz et al., 2010). Another example comes from a recent study of the T cell receptor-regulated phosphoproteome (Navarro et al., 2011). Navarro and colleagues identified a subset of transcriptional regulators as significantly enriched among the phosphoproteins measured in their T cell model system. Among these regulators, HDAC7

appeared as a key regulator of the expression of the IL-2 receptor by T cells (Navarro et al., 2011).

Besides the numerous studies of PTMs, measurements of total protein abundance across various cell types have led to important advances in our knowledge of the division of labor of immune cell subsets. For example, measurements of protein levels in classical and plasmacytoid DCs from the mouse spleen have uncovered differences in pathogen sensor expression among these cells (Luber et al., 2010). Altogether, the studies described above as examples clearly demonstrate the power of MS-based proteomics to map intracellular processes such as signal transduction.

Integrative Analyses. Building upon these developments in dissecting various regulatory layers of cell biology, integrative studies that aim to tackle the challenge of dissecting immune cellular circuitries have emerged. Broadly speaking, they focus on explaining immunological functions by providing functional and physical information across various regulatory layers. For example, a landmark study on the TNF-α/NF-κB signal transduction pathway has paved the way to these integrative studies (Bouwmeester et al., 2004). In this study, the authors focused on 32 known signaling and transcriptional components of the tumor necrosis factor (TNF)-α pathway. Then, they placed each of these components within their functional context in the activated pathway by identifying their binding partners using MS-based proteomics. Furthermore, functional perturbations of selected binding partners using RNAi led to the identification of novel components of the TNF-α pathways such as the E3 ubiquitin ligase TRAF7. This study constitutes one of the first examples of an integrated approach to dissect a signal transduction pathway from both physical and functional standpoints.

More recent works relying on a similar conceptual framework have also participated in our systems-level understanding of processes important to DC biology such as antigen presentation by major histocompatibility complex (MHC) molecules (Neefjes et al., 2011; Paul et al., 2011), and pathogen-sensing pathways (Chiang et al., 2012; Li et al., 2011). To further our knowledge of MHC class II biology, one of the studies mentioned above (Paul et al., 2011) has taken advantage of RNAi perturbations to screen for new modulators of MHC II trafficking and surface expression in a cell line model. Using a genome-wide RNAi screen, a couple hundred of candidate regulators were shown to have an impact on MHC II expression at the cell surface and/or MHC II peptide loading (Paul et al., 2011). Several GTPases and motor proteins not previously linked to the MHC II pathway were identified as regulators of the transport of MHC II within antigen-presenting cells (APCs) (Paul et al., 2011). These types of studies will help in increasing the pace of discovery of regulators of MHC processes, and point towards clear mechanistic hypotheses that may be tested by standard biochemical and cell biological approaches. Ultimately, given the central role of MHC presentation in normal and pathological immunological events, such detailed understanding of antigen processing may lead to improved therapeutics in the future (Neefjes et al., 2011).

In the field of pathogen sensing, recent studies (Chiang et al., 2012; Li et al., 2011) have followed approaches conceptually similar to the works mentioned above. In the study reported by Li and colleagues, 58 components of the signaling pathways leading to type I IFN production were subjected to immunoprecipitation followed by MS to identify binding partners (Li et al., 2011). Among the numerous interacting proteins found for these 58 known components, MIB1 and MIB2 were found to play a redundant role in regulating type I IFN production via the TBK1-IRF3 signaling axis. MIBs were shown to directly ubiquitinylate TBK1, which was important for TBK1 activity and downstream antiviral gene expression in the context of viral infection or stimulation with viral ligands (Li et al., 2011). On the other hand, Chiang and colleagues

performed a genome-wide RNAi screen to identify regulators of TLR7- and TLR9-mediated activation of NF-κB (Chiang et al., 2012). Using a variety of independent evidence of functionality together with re-screening with multiple siRNAs, they were able to identify nearly 200 regulators of these pathways with putative roles at various levels of regulation such as transcription, signaling, or TLR intracellular trafficking (Chiang et al., 2012). For example, HRS, a component of the endosomal sorting complex required for transport (ESCRT) machinery, was shown to play a role in the delivery of TLR7 and 9 to endolysosomes from where these receptors have been shown to initiate signal transduction (Chiang et al., 2012).

Overall, the selected studies described in this section clearly exemplify the growing role of systems approaches in helping to understand immunological processes, and to generate hypotheses and models regarding these processes. Many challenges remain, including the use of immune cell types that are more relevant to the pathways and processes under study. For example, genome-wide RNAi screen and targeted interaction proteomic studies were performed in cancer cell lines that may not express all of the components relevant to the pathways under study. Furthermore, there is a need for systematic approaches that start addressing mechanistic questions – a challenge which is not limited to immunology. For example, how can we better define the function of candidate emanating from large-scale RNAi screens?

1.4. Pathogen-Sensing Pathways of Innate Immunity

Toll-Like Receptors: Overview and Significance. The past decade has witnessed fundamental progress in our understanding of the innate immune system – the first line of host defense against pathogens. The identification of host sensors and their cognate ligands has shed light on the initiation of innate immunity. These advances stem from a body of theoretical and experimental work that led to the identification of sensors for microbial molecules in

Drosophila (Toll protein), and subsequently in mammals (Toll-like receptor; TLR) (Beutler and Rietschel, 2003; Kawai et al., 1999; Lemaitre et al., 1996; Medzhitov, 2009; Medzhitov et al., 1997; Poltorak et al., 1998). Remarkably, these sensors are evolutionary conserved in vertebrates and insects, and share structural and mechanistic features with immune sensors of infection in plants (Ronald and Beutler, 2010).

To date, several families of innate immune sensors have been revealed across most intracellular compartments: transmembrane TLRs and C-type lectin receptors (CLRs), and cytosolic nucleotide-binding oligomerization domain (NOD)-like receptors (NLRs) and retinoic acid-inducible gene (RIG)-I-like receptors (RLRs) (Takeuchi and Akira, 2010). Even though many questions remain regarding the biology of TLRs, they are currently considered to be the best-characterized sensors of microorganism- and virus-derived molecules in mammals (Medzhitov et al., 2011). Ten human and twelve mouse TLRs have been uncovered, and exogenous and/or endogenous ligands have been identified for most of them. For example, TLR2 forms heterodimers with TLR1 and TLR6 to bind bacterial lipoproteins; TLR4 recognizes lipopolysaccharide together with MD-2 and CD14; and TLR3, 7, 8 and 9 detects nucleic acids (Takeuchi and Akira, 2010).

TLRs have been implicated in a variety of infectious and non-infectious diseases, ranging from cancer to autoimmunity. Strong genetic evidence for these implications has emerged in recent years. For example, mouse and human genetic studies have associated TLRs and downstream pathway components with infectious disease susceptibility and progression (Beutler et al., 2006; Casanova et al., 2011; Chapman and Hill, 2012; Takeuchi and Akira, 2010). Genome-wide association studies (GWAS) have participated in connecting the TLR pathways to increased susceptibility to a variety of autoimmune disorders, such as systemic lupus erythematosus (SLE) or rheumatoid arthritis (Xavier and Rioux, 2008).

This broad range of biological functions and dysfunctions makes TLRs attractive targets for the development of immunotherapies in many clinical contexts (Hennessy et al., 2010; Kanzler et al., 2007). For example, in infectious diseases, the use of TLR agonists as vaccine adjuvants is an area of intense investigation. The effects of existing adjuvants have also been shown to rely on signals from TLRs and other innate immune sensors (Hennessy et al., 2010; Ishii and Akira, 2007). Furthermore, TLRs and other innate immune sensing pathways play important roles in tumorigenesis (Grivennikov et al., 2010), and in other chronic inflammatory disorders such as autoimmunity (Mills, 2011). Altogether, these findings have spurred many clinical programs to translate the knowledge of TLR biology into therapeutics (Hennessy et al., 2010; Kanzler et al., 2007), which illustrate the important need for a precise dissection of the TLR system across regulatory layers and cell types.

Control of Dendritic Cell Processes by TLRs. TLRs act as sensors that allow the host to perceive the presence of microorganisms, an ongoing infection, or tissue damage. The signals emanating from TLRs are not only key in the initiation of innate immune responses, but also in the mounting of adaptive immune responses (Iwasaki and Medzhitov, 2010; Paul, 2011). DCs are perhaps the most important cellular nexus that integrate signals coming from TLRs and other innate immune receptors, and relay them to downstream adaptive immune cells such as T lymphocytes (Steinman and Banchereau, 2007).

DCs consist of multiple subtypes that are collectively present in nearly every organ of the body (Idoyaga and Steinman, 2011). Upon detection of microbial constituents and/or other environmental signals, DCs ingest and process antigens for presentation at their cell surface on major histocompatibility complex (MHC) molecules, and migrate from the tissue into the draining lymph node where priming of antigen-specific lymphocytes occurs. Recent years have led to

major advances in understanding DC biology in terms, for example, of their developmental pathways (Liu and Nussenzweig, 2010) or their roles in humans (Collin et al., 2011). Because the primary focus of this thesis work is to decipher intracellular circuits, we will focus on introducing the initial changes induced in DCs by TLR signals (Watts et al., 2010), together with some of the associated regulatory mechanisms of TLR function.

TLR Signaling. Upon TLR engagement, DCs activate multiple pathways leading to an increase in: (1) antigen uptake; (2) MHC molecule synthesis, trafficking, and turnover; (3) motility and migration; and (4) transcriptional and translational activity, including production of soluble mediators such as pro-inflammatory and antiviral cytokines (Watts et al., 2010). All of these changes are controlled, at least in part, by the cascades of signals acting directly downstream of TLRs. TLRs are type I transmembrane proteins composed of an ectodomain containing leucine-rich repeats (LRRs), a single transmembrane domain, and a cytoplasmic Toll/IL-1 receptor (TIR) domain (Takeuchi and Akira, 2010). Upon ligand binding, TLRs form homo- or hetero-dimers, which leads to the recruitment of cytosolic adaptor molecules and initiates signaling. LRRs of the TLR ectodomains bind to cognate ligand, which is thought to trigger conformational changes responsible for the intracellular recruitment of TIR-domain containing adaptors. The association between these TIR-domain adaptors and the intracellular TIR domain of TLRs constitute the first step in the signaling cascade. The selective recruitment of subsets of TIR adaptors helps explaining, at least in part, the specificity in TLR signaling.

Five TIR domain-containing adaptors have been reported to be functional in the TLR pathways: myeloid differentiation primary-response protein 88 (MYD88), TIR-domain-containing adaptor inducing IFNβ (TRIF), TRIF-related adaptor molecule (TRAM), TIRAP that is also known as Mal, and Sterile-alpha and Armadillo motif-containing protein (SARM) (Takeuchi and Akira, 2010). Classically, the TLR signaling cascades have been subdivided into two pathways depending on

the use of the TIR domain-containing adaptors MYD88 and TRIF. These two pathways are referred to as the MYD88-dependent and the TRIF-dependent pathways, respectively. All TLRs but TLR3 have been shown to recruit the MYD88 to their cytoplasmic TIR domain. Some TLRs, such as TLR7 and 9, rely completely on MYD88 for signal transduction, whereas TLR3 depends on the recruitment of TRIF for signaling. TLR4 is the only TLR that signal through both MYD88 and TRIF.

The MYD88-dependent pathway initiates with the association of death domains (DD) between MYD88 and IL-1R-associated kinases (IRAKs), a family of serine-threonine protein kinases. MYD88 has been shown to first bind IRAK4, which in turn activates IRAK1 and 2 (Takeuchi and Akira, 2010). IRAKs activate the E3 ubiquitin ligase TNFR-associated factor 6 (TRAF6), which in turn, in combination with the E2 ubiquitin-conjugating enzymes Ubc13 and Uev1A, activates a complex comprised of TGF-β-activated kinase 1 (TAK1), TAK1-binding protein 1 (TAB1), TAB2, and TAB3. The formation of the TAK1 complex leads to the activation of both mitogen-activated protein kinases (MAPKs) and NF-κB pathways. The first members of the MAPK pathway to be activated by the TAK1 complex are MAPK kinase (MAPKK) 3 and 6, leading to the subsequent activation of the three canonical MAPK pathways: Jun kinase (JNK), p38, and ERK. Ultimately, MAPKs activate downstream transcriptional regulators that include CREB and AP-1. In addition, the TAK1 complex phosphorylates IκB kinase (IKK)-β, which is part of the IKK complex comprising IKK-α, IKK-β, and NF-κB essential modulator (NEMO). Upon activation, the IKK complex phosphorylates the NF-κB inhibitory protein IκBα, leading to its degradation via the ubiquitin pathway to the proteasome. Upon degradation of IκBα, NF-κB is relieved from cytoplasmic sequestration and translocates to the nucleus, where it will participate in the regulation of pro-inflammatory gene production, such as tumor necrosis factor (TNF), interleukin (IL)-6, IL-12, or IL-1. Interestingly, the MYD88-dependent pathway varies depending on the cell-type considered. Whereas the canonical pathways described above are known to occur in

classical DCs (cDCs), in plasmacytoid DCs (pDCs) stimulated with TLR7 or 9 agonists, MYD88 forms a complex containing TRAF6 and 3, IRAK family members, IKK-α and IRF7, among others, which leads to the phosphorylation of IRF7 followed by its translocation into the nucleus. Nuclear IRF7 controls the expression of type I interferons (IFNs). In addition, this MYD88-dependent multiprotein complex downstream of TLR7 and 9 also controls the activation of the NF-κB and MAPK pathways.

The TRIF-dependent pathway occurs downstream of only TLR4 and 3 (Takeuchi and Akira, 2010). While TLR3 recruits TRIF directly to its intracellular TIR domain, TLR4 relies on an intermediate TIR domain-containing adaptor, namely TRAM, for TRIF recruitment and activation. TRIF harbors TRAF-binding motifs allowing the formation of a complex between TRIF and TRAF3 and 6. In addition, TRIF associates with receptor-interacting proteins 1 (RIP1) and 3 (RIP3) through a RIP homotypic interaction motif (RHIM). This TRIF complex has been described to activate two pathways leading to the transcriptional activation of both pro-inflammatory genes and type I IFNs. For pro-inflammatory gene activation, the TRIF complex activates the NF-κB and MAPK pathways via TRAF6, RIP1, TNFR-associated death domain protein (TRADD), and TAK1. TRADD, upon association with FAS-associated death domain-containing protein (FADD), ubiquitinates RIP1, which in turn leads to the activation of NF-κB and MAPKs to induce pro-inflammatory gene expression. For type I IFN induction, the TRIF complex activates via the ubiquitin ligase TRAF3 two IKK-related kinases: TANK-binding kinase 1 (TBK1) and IKK-i (also known as IKK-ε). TBK1 and IKK-i phosphorylate IRF3 and 7, which, upon hetero- and homo-dimerization, migrate into the nucleus to induce the expression of type I IFNs and IFN-inducible genes. Furthermore, TBK1 and IKK-i are modulated by a variety of factors such as TRAF family member-associated NF-κB activator (TANK), NAK-associated protein 1 (NAP1), and the TBK1 adaptor (SINTBAD) (Takeuchi and Akira, 2010). The modulation of TBK1 and IKK-i by these factors is motif-dependent. Indeed, these factors contain a TBK1-binding

motif and exhibit similarities in their coiled coil domains. Importantly, the relationships between these various factors still remain to be elucidated.

Overall, the canonical pathways downstream of MYD88 include MAPKs and NF-κB, which drives the induction of pro-inflammatory cytokines, whereas the TRIF-dependent pathway culminates with the activation of NF-κB via TRAF6 and TAK1, and of IRF3 via TBK1 and IKKi, thus leading to the induction of both pro-inflammatory cytokines and type I interferons (IFNs) (Kawai and Akira, 2010). The balance between these pathways is critical in the control of infections, and their dysregulation has been associated with human disorders, including cancer and autoimmunity.

Multilayer Regulation of the TLR System. TLR functions are regulated by control mechanisms operating at various cellular and molecular levels, which largely contribute to shaping TLR signaling and transcriptional networks. For example, TLRs may rely on different downstream components depending on the DC subset in which they are expressed (Blasius and Beutler, 2010). Indeed, while TLR7 or TLR9 activation on pDCs induces both inflammatory cytokines and type I IFNs, it only leads to inflammatory cytokine production in cDCs. TLR function is also tightly regulated by a variety of accessory molecules such as mediators of ligand delivery and/or recognition (*e.g.*, HMGB proteins), chaperones, molecules facilitating intracellular trafficking or recycling of TLRs (*e.g.*, Unc93b1), and TLR-processing enzymes (Lee et al., 2012). Some of these cofactors ensure that TLRs are delivered to specific subcellular compartments (*e.g.*, endocytic pathway, plasma membrane), thus controlling the types of ligands to which given TLRs are exposed to, and the access to assigned signaling components and platforms (Barton and Kagan, 2009).

Additional control checkpoints operate at the transcriptional and post-transcriptional levels, which also participate in inducing and shutting off TLR-induced signals and downstream effectors. For example, the transcription factors (TFs) acting downstream of TLR4, which recognizes LPS, have been classified into three classes: (*i*) constitutively expressed but latent (*e.g.*, NF-κB, IRF3), (*ii*) synthesized de novo upon TLR induction (*e.g.*, Cebpd), and (*iii*) induced during cell differentiation (*e.g.*, Cebpb, Runx1) (Medzhitov and Horng, 2009). Altogether, these various classes of TFs contribute to various phases of the TLR transcriptional response. While the rapid induction of these transcriptional changes is vital to the host, shutting them off and reducing the levels of inflammatory cytokines is equally as important. For example, specific RNA-binding proteins are important in regulating the stability or degradation of inflammatory cytokine mRNAs (Anderson, 2010).

1.5. Main Goals and Questions

Altogether, the regulatory processes highlighted above provide a brief description of some of the strategies contributing to the control of TLR function. The main goal of this thesis work is to start tackling the challenge of dissecting the modules and organizational rules driving the 'signaling-to-transcription' TLR network. To do so, we designed experimental and analytic pipelines – easily transposable to other systems – and applied them to the study of TLRs in DCs. In brief, we build upon an initial approach focused on deciphering the transcriptional regulatory layer of TLRs (**Chapter 2**), by expanding it to signaling regulatory layers (**Chapter 3**), and by moving towards an integrative analysis of functional and physical interactions within and between these layers (**Chapter 4**).

Some of the biological and immunological questions that have driven this work include: (*i*) how are pathogen-specific responses shaped by the regulatory circuitries of the TLR system; (*ii*)

what are the missing components and modules that account for the responses and help explain their overall functioning; (*iii*) how to address these questions in a large-scale and unbiased manner; and (*iv*) how to construct a model network, formulate hypotheses from it, and manipulate it?

1.6. References

Anderson, P. (2010). Post-transcriptional regulons coordinate the initiation and resolution of inflammation. Nature reviews Immunology *10*, 24-35.

Barton, G.M., and Kagan, J.C. (2009). A cell biological view of Toll-like receptor function: regulation through compartmentalization. Nature reviews Immunology *9*, 535-542.

Benoist, C., Germain, R.N., and Mathis, D. (2006). A plaidoyer for 'systems immunology'. Immunological reviews *210*, 229-234.

Berry, M.P., Graham, C.M., McNab, F.W., Xu, Z., Bloch, S.A., Oni, T., Wilkinson, K.A., Banchereau, R., Skinner, J., Wilkinson, R.J.*, et al.* (2010). An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. Nature *466*, 973-977.

Beutler, B., Jiang, Z., Georgel, P., Crozat, K., Croker, B., Rutschmann, S., Du, X., and Hoebe, K. (2006). Genetic analysis of host resistance: Toll-like receptor signaling and immunity at large. Annual review of immunology *24*, 353-389.

Beutler, B., and Rietschel, E.T. (2003). Innate immune sensing and its roots: the story of endotoxin. Nature reviews Immunology *3*, 169-176.

Blasius, A.L., and Beutler, B. (2010). Intracellular toll-like receptors. Immunity *32*, 305-315.

Bouwmeester, T., Bauch, A., Ruffner, H., Angrand, P.O., Bergamini, G., Croughton, K., Cruciat, C., Eberhard, D., Gagneur, J., Ghidelli, S.*, et al.* (2004). A physical and functional map of the human TNF-alpha/NF-kappa B signal transduction pathway. Nature cell biology *6*, 97-105.

Casanova, J.L., Abel, L., and Quintana-Murci, L. (2011). Human TLRs and IL-1Rs in host defense: natural insights from evolutionary, epidemiological, and clinical genetics. Annual review of immunology *29*, 447-491.

Chapman, S.J., and Hill, A.V. (2012). Human genetic susceptibility to infectious disease. Nature reviews Genetics *13*, 175-188.

Chiang, C.Y., Engel, A., Opaluch, A.M., Ramos, I., Maestre, A.M., Secundino, I., De Jesus, P.D., Nguyen, Q.T., Welch, G., Bonamy, G.M.*, et al.* (2012). Cofactors Required for TLR7- and TLR9-Dependent Innate Immune Responses. Cell host & microbe *11*, 306-318.

Choudhary, C., and Mann, M. (2010). Decoding signalling networks by mass spectrometry-based proteomics. Nature reviews Molecular cell biology *11*, 427-439.

Collin, M., Bigley, V., Haniffa, M., and Hambleton, S. (2011). Human dendritic cell deficiency: the missing ID? Nature reviews Immunology *11*, 575-583.

Costanzo, M., Baryshnikova, A., Myers, C.L., Andrews, B., and Boone, C. (2011). Charting the genetic interaction map of a cell. Current opinion in biotechnology *22*, 66-74.

Fraser, I.D., and Germain, R.N. (2009). Navigating the network: signaling cross-talk in hematopoietic cells. Nature immunology *10*, 327-331.

Germain, R.N., Meier-Schellersheim, M., Nita-Lazar, A., and Fraser, I.D. (2011). Systems biology in immunology: a computational modeling perspective. Annual review of immunology *29*, 527-585.

Gilchrist, M., Thorsson, V., Li, B., Rust, A.G., Korb, M., Roach, J.C., Kennedy, K., Hai, T., Bolouri, H., and Aderem, A. (2006). Systems biology approaches identify ATF3 as a negative regulator of Toll-like receptor 4. Nature *441*, 173-178.

Grivennikov, S.I., Greten, F.R., and Karin, M. (2010). Immunity, inflammation, and cancer. Cell *140*, 883-899.

Guiducci, C., Gong, M., Xu, Z., Gill, M., Chaussabel, D., Meeker, T., Chan, J.H., Wright, T., Punaro, M., Bolland, S.*, et al.* (2010). TLR recognition of self nucleic acids hampers glucocorticoid activity in lupus. Nature *465*, 937-941.

Hartwell, L.H., Hopfield, J.J., Leibler, S., and Murray, A.W. (1999). From molecular to modular cell biology. Nature *402*, C47-52.

Heng, T.S., and Painter, M.W. (2008). The Immunological Genome Project: networks of gene expression in immune cells. Nature immunology *9*, 1091-1094.

Hennessy, E.J., Parker, A.E., and O'Neill, L.A. (2010). Targeting Toll-like receptors: emerging therapeutics? Nature reviews Drug discovery *9*, 293-307.

Ideker, T., Galitski, T., and Hood, L. (2001). A new approach to decoding life: systems biology. Annual review of genomics and human genetics *2*, 343-372.

Idoyaga, J., and Steinman, R.M. (2011). SnapShot: Dendritic Cells. Cell *146*, 660-660 e662.

Ishii, K.J., and Akira, S. (2007). Toll or toll-free adjuvant path toward the optimal vaccine development. Journal of clinical immunology *27*, 363-371.

Iwasaki, A., and Medzhitov, R. (2010). Regulation of adaptive immunity by the innate immune system. Science *327*, 291-295.

Jenner, R.G., and Young, R.A. (2005). Insights into host responses against pathogens from transcriptional profiling. Nature reviews Microbiology *3*, 281-294.

Kanzler, H., Barrat, F.J., Hessel, E.M., and Coffman, R.L. (2007). Therapeutic targeting of innate immunity with Toll-like receptor agonists and antagonists. Nature medicine *13*, 552-559.

Kawai, T., Adachi, O., Ogawa, T., Takeda, K., and Akira, S. (1999). Unresponsiveness of MyD88-deficient mice to endotoxin. Immunity *11*, 115-122.

Kawai, T., and Akira, S. (2010). The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. Nature immunology *11*, 373-384.

Kitano, H. (2002). Computational systems biology. Nature *420*, 206-210.

Lander, E.S. (2011). Initial impact of the sequencing of the human genome. Nature *470*, 187-197.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W.*, et al.* (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860-921.

Lee, C.C., Avalos, A.M., and Ploegh, H.L. (2012). Accessory molecules for Toll-like receptors and their function. Nature reviews Immunology *12*, 168-179.

Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J.M., and Hoffmann, J.A. (1996). The dorsoventral regulatory gene cassette spatzle/Toll/cactus controls the potent antifungal response in Drosophila adults. Cell *86*, 973-983.

Li, S., Wang, L., Berman, M., Kong, Y.Y., and Dorf, M.E. (2011). Mapping a dynamic innate immunity protein interaction network regulating type I interferon production. Immunity *35*, 426-440.

Litvak, V., Ramsey, S.A., Rust, A.G., Zak, D.E., Kennedy, K.A., Lampano, A.E., Nykter, M., Shmulevich, I., and Aderem, A. (2009). Function of C/EBPdelta in a regulatory circuit that discriminates between transient and persistent TLR4-induced signals. Nature immunology *10*, 437-443.

Liu, K., and Nussenzweig, M.C. (2010). Origin and development of dendritic cells. Immunological reviews *234*, 45-54.

Luber, C.A., Cox, J., Lauterbach, H., Fancke, B., Selbach, M., Tschopp, J., Akira, S., Wiegand, M., Hochrein, H., O'Keeffe, M.*, et al.* (2010). Quantitative proteomics reveals subset-specific viral recognition in dendritic cells. Immunity *32*, 279-289.

MacArthur, D.G., Balasubramanian, S., Frankish, A., Huang, N., Morris, J., Walter, K., Jostins, L., Habegger, L., Pickrell, J.K., Montgomery, S.B.*, et al.* (2012). A systematic survey of loss-of-function variants in human protein-coding genes. Science *335*, 823-828.

Medzhitov, R. (2009). Approaching the asymptote: 20 years later. Immunity *30*, 766-775.

Medzhitov, R., and Horng, T. (2009). Transcriptional control of the inflammatory response. Nature reviews Immunology *9*, 692-703.

Medzhitov, R., Preston-Hurlburt, P., and Janeway, C.A., Jr. (1997). A human homologue of the Drosophila Toll protein signals activation of adaptive immunity. Nature *388*, 394-397.

Medzhitov, R., Shevach, E.M., Trinchieri, G., Mellor, A.L., Munn, D.H., Gordon, S., Libby, P., Hansson, G.K., Shortman, K., Dong, C.*, et al.* (2011). Highlights of 10 years of immunology in Nature Reviews Immunology. Nature reviews Immunology *11*, 693-702.

Mills, K.H. (2011). TLR-dependent T cell activation in autoimmunity. Nature reviews Immunology *11*, 807-822.

Nakaya, H.I., Wrammert, J., Lee, E.K., Racioppi, L., Marie-Kunze, S., Haining, W.N., Means, A.R., Kasturi, S.P., Khan, N., Li, G.M.*, et al.* (2011). Systems biology of vaccination for seasonal influenza in humans. Nature immunology *12*, 786-795.

Navarro, M.N., Goebel, J., Feijoo-Carnero, C., Morrice, N., and Cantrell, D.A. (2011). Phosphoproteomic analysis reveals an intrinsic pathway for the regulation of histone deacetylase 7 that controls the function of cytotoxic T lymphocytes. Nature immunology *12*, 352-361.

Neefjes, J., Jongsma, M.L., Paul, P., and Bakke, O. (2011). Towards a systems understanding of MHC class I and MHC class II antigen presentation. Nature reviews Immunology *11*, 823-836.

Northrup, D.L., and Zhao, K. (2011). Application of ChIP-Seq and related techniques to the study of immune function. Immunity *34*, 830-842.

Novershtern, N., Subramanian, A., Lawton, L.N., Mak, R.H., Haining, W.N., McConkey, M.E., Habib, N., Yosef, N., Chang, C.Y., Shay, T.*, et al.* (2011). Densely interconnected transcriptional circuits control cell states in human hematopoiesis. Cell *144*, 296-309.

Ochman, H., and Raghavan, R. (2009). Systems biology. Excavating the functional landscape of bacterial cells. Science *326*, 1200-1201.

Pascual, V., Chaussabel, D., and Banchereau, J. (2010). A genomic approach to human autoimmune diseases. Annual review of immunology *28*, 535-571.

Paul, P., van den Hoorn, T., Jongsma, M.L., Bakker, M.J., Hengeveld, R., Janssen, L., Cresswell, P., Egan, D.A., van Ham, M., Ten Brinke, A.*, et al.* (2011). A Genome-wide multidimensional RNAi screen reveals pathways controlling MHC class II antigen presentation. Cell *145*, 268-283.

Paul, W.E. (2011). Bridging innate and adaptive immunity. Cell *147*, 1212-1215.

Poltorak, A., He, X., Smirnova, I., Liu, M.Y., Van Huffel, C., Du, X., Birdwell, D., Alejos, E., Silva, M., Galanos, C.*, et al.* (1998). Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. Science *282*, 2085-2088.

Pulendran, B., Li, S., and Nakaya, H.I. (2010). Systems vaccinology. Immunity *33*, 516-529.

Querec, T.D., Akondy, R.S., Lee, E.K., Cao, W., Nakaya, H.I., Teuwen, D., Pirani, A., Gernert, K., Deng, J., Marzolf, B.*, et al.* (2009). Systems biology approach predicts immunogenicity of the yellow fever vaccine in humans. Nature immunology *10*, 116-125.

Ronald, P.C., and Beutler, B. (2010). Plant and animal sensors of conserved microbial signatures. Science *330*, 1061-1064.

Ross, J. (2008). From the determination of complex reaction mechanisms to systems biology. Annual review of biochemistry *77*, 479-494.

Schena, M., Shalon, D., Davis, R.W., and Brown, P.O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science *270*, 467-470.
Steinman, R.M., and Banchereau, J. (2007). Taking dendritic cells into medicine. Nature *449*, 419-426.

Suzuki, H., Forrest, A.R., van Nimwegen, E., Daub, C.O., Balwierz, P.J., Irvine, K.M., Lassmann, T., Ravasi, T., Hasegawa, Y., de Hoon, M.J.*, et al.* (2009). The transcriptional network that controls growth arrest and differentiation in a human myeloid leukemia cell line. Nature genetics *41*, 553-562.

Takeuchi, O., and Akira, S. (2010). Pattern recognition receptors and inflammation. Cell *140*, 805-820.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A.*, et al.* (2001). The sequence of the human genome. Science *291*, 1304-1351.

Watts, C., West, M.A., and Zaru, R. (2010). TLR signalling regulated antigen presentation in dendritic cells. Current opinion in immunology *22*, 124-130.

Weintz, G., Olsen, J.V., Fruhauf, K., Niedzielska, M., Amit, I., Jantsch, J., Mages, J., Frech, C., Dolken, L., Mann, M.*, et al.* (2010). The phosphoproteome of toll-like receptor-activated macrophages. Molecular systems biology *6*, 371.

Xavier, R.J., and Rioux, J.D. (2008). Genome-wide association studies: a new window into immune-mediated diseases. Nature reviews Immunology *8*, 631-643.

Zak, D.E., and Aderem, A. (2009). Systems biology of innate immunity. Immunological reviews *227*, 264-282.

Chapter 2 – Unbiased Reconstruction of the Mammalian Transcriptional Network Mediating TLR Responses

2.1. Author Contributions

Reference (reproduced with permission from *Science*)**:**

Amit, I., Garber, M.*, Chevrier, N.*, Leite, A.P.*, Donner, Y.*, Eisenhaure, T., Guttman, M., Grenier, J.K., Li, W., Zuk, O., Schubert, L.A., Birditt, B., Shay, T., Goren, A., Zhang, X., Smith, Z., Deering, R., McDonald, R.C., Cabili, M., Bernstein, B.E., Rinn, J.L., Meissner, A., Root, R.E., Hacohen, N.*, Regev, A.* (2009). Unbiased reconstruction of a mammalian transcriptional network mediating pathogen responses. Science *326*, 257-263. (* equal contributions)

- Ido Amit conceived the study, designed and performed experiments, analyzed and interpreted data, and wrote the manuscript.
- Manuel Garber performed computational analysis of the NanoString dataset and network modeling.
- Nicolas Chevrier participated in designing and performing most experiments (except ChIP).
- Ana Paula Leite performed comparative analysis of our microarray datasets with published work.
- Yoni Donner created the GeneSelector algorithm.
- Thomas Eisenhaure assisted with microarray and RNAi experiments.
- Mitchell Guttman assisted with computational analysis.
- Jennifer K. Grenier assisted with RNAi experimental design.

- Weibo Li assisted with shRNA cloning.
- Or Zuk assisted with transcription factor binding site analysis.
- Lisa A. Schubert and Brian Birditt assisted with operating NanoString equipments.
- Tal Shay assisted with computational analysis.
- Alon Goren, Xiaolan Zhang and Zachary Smith assisted with ChIP experiments.
- Raquel Deering and Rebecca C. McDonald assisted with RNAi experimental design.
- Moran Cabili assisted with analysis of disease-associated genes.
- Bradley E. Bernstein, John L. Rinn, and Alex Meissner supervised ChIP experiments.
- David E. Root supervised RNAi experiments and provided reagents.
- Nir Hacohen supervised the project and manuscript writing.
- Aviv Regev supervised the project and manuscript writing.

2.2. Introduction

Regulatory networks controlling gene expression serve as decision-making circuits within cells. For example, when immune dendritic cells are exposed to viruses, bacteria or fungi, they respond with transcriptional programs that are specific to each pathogen (Huang et al., 2001) and are essential for establishing appropriate immunological outcomes (Kawai and Akira, 2009). These responses are initiated through specific receptors, such as Toll-like receptors (TLRs), that distinguish broad pathogen classes, and are propagated through well-characterized signaling cascades (Kawai and Akira, 2009). However, little is known about how the transcriptional network is wired to produce specific outputs.

Two major observational strategies have associated regulators with their putative targets on a genome scale (Kim et al., 2009): *cis*-regulatory models rely on the presence of predicted transcription factor binding sites in the promoters of target genes (Kim et al., 2009; Ramsey et al., 2008; Suzuki et al., 2009), whereas *trans*-regulatory models are based on correlations

between regulator and target expression (Kim et al., 2009; Ramsey et al., 2008; Suzuki et al., 2009; Segal et al., 2003). Because promoter binding sites and correlated expression are weak predictors of functional regulator-target linkages, such approaches are limited in their ability to produce reliable models of transcriptional networks (Kim et al., 2009). A complementary strategy is to systematically perturb every regulatory input and measure its effect on expression of gene targets. This strategy has been successfully employed in yeast (Capaldi et al., 2008; Hu et al., 2007; Workman et al., 2006) and sea urchin (Erwin and Davidson, 2009), but not in mammals.

2.3. A Perturbation-Based Strategy for Network Reconstruction

We developed a perturbation strategy for reconstructing transcriptional networks in mammalian cells, and applied it to determine a network controlling the responses of DCs to pathogens (**Figure 2.1**).

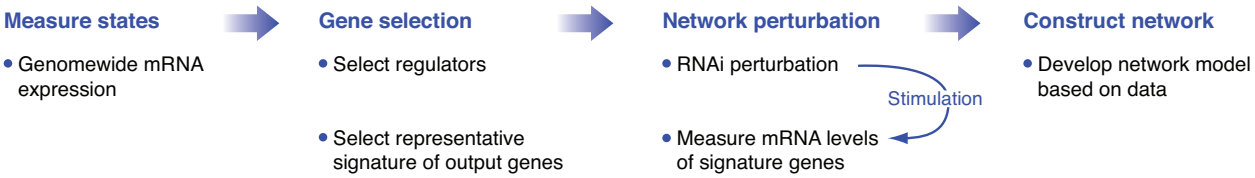

Figure 2.1. A Systematic Strategy for Network Reconstruction
The strategy consists of four steps (left to right): state measurement using arrays; selection of regulators and response signatures; network perturbation with shRNAs against each regulator, followed by measurement of signature genes; and network reconstruction from the perturbational data.

First, we profiled gene expression at nine time points following stimulation with five pathogen-derived components and identified specific and shared genes that respond to each stimulus**.** We used these profiles to identify 144 candidate regulators whose expression changed in response to at least one stimulus. We also identified a signature of 118 marker genes that captures the complexity of the response. We generated a validated lentiviral shRNA library for 125 of the 144

candidate regulators, used it to systematically perturb each of the regulators in DCs, stimulated the cells with a pathogen component, and profiled the expression of the 118 gene signature (Geiss et al., 2008). Finally, we used the measurements from the perturbed cells to derive a validated model of the regulatory network.

2.4. Gene Expression Programs in Response to TLR Agonists

We measured genome-wide expression profiles in DCs exposed to Pam3CSK4 (PAM), a synthetic mimic of bacterial lipopeptides; polyinosinic:polycytidylic acid (poly(I:C)), a viral-like double-stranded RNA; lipopolysaccharide (LPS), a purified component from Gram negative *Escherichia coli*; gardiquimod, a small molecule agonist; and CpG, a synthetic single-stranded DNA. These compounds are known agonists of TLR2, TLR3, TLR4, TLR7 and TLR9, respectively. poly(I:C) also activates MDA-5, and LPS can also act through co-receptors such as CD14. We therefore refer to the ligands rather than their receptors for clarity. Based on pilot experiments (Amit et al., 2009), we measured mRNA expression at 0.5, 1, 2, 4, 6, 8, 12, 16, and 24 hours following stimulation with these pathogen components (**Figure 2.2A**).

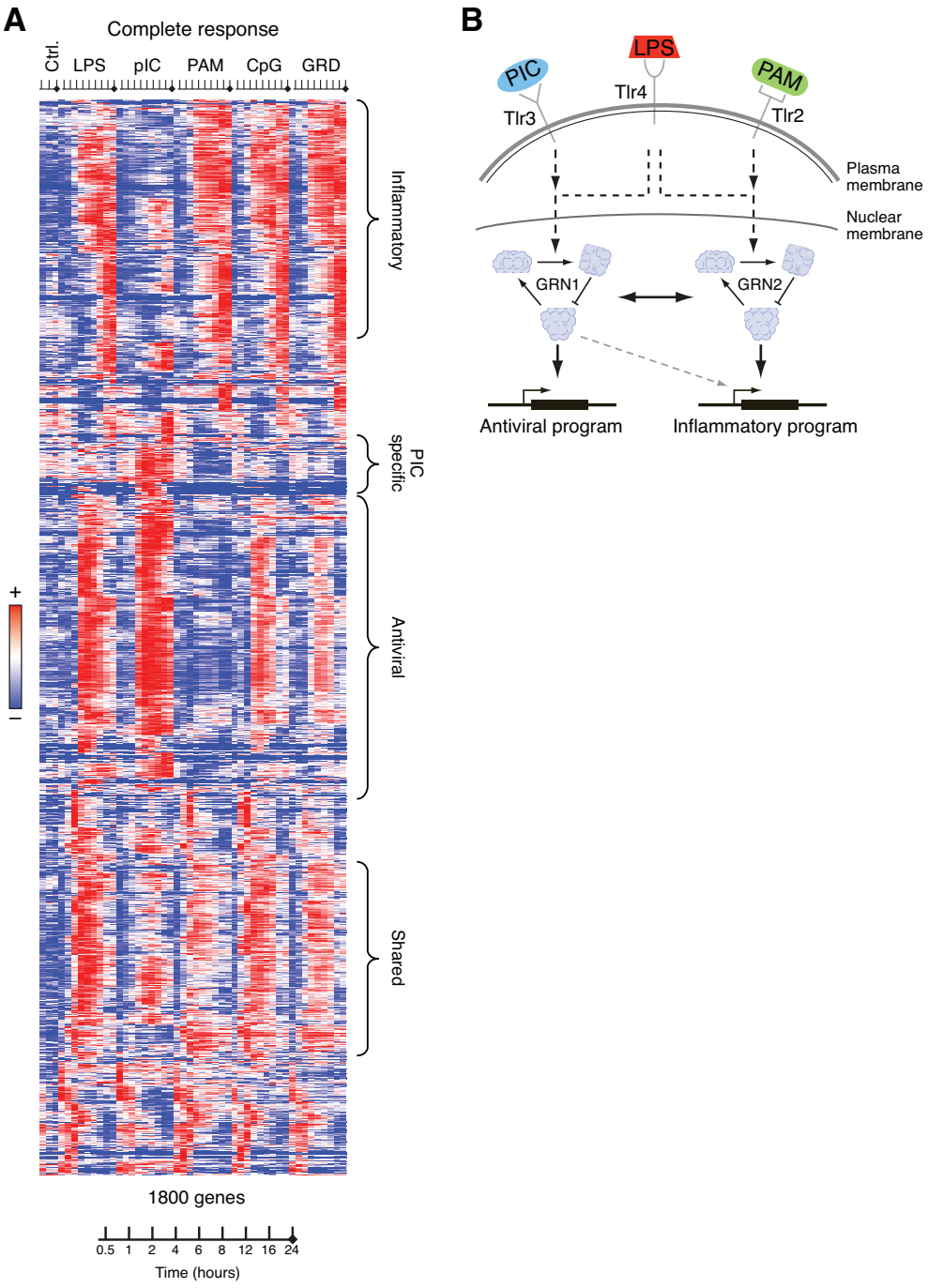

Figure 2.2. Gene Expression Response to Pathogen Components

(A) mRNA profiles of the 1800 genes whose expression was induced by a factor of at least 1.7 from baseline level in both duplicates of at least one time point in CD11c+ DCs stimulated with the indicated pathogen component across a time course of 0, 0.5, 1, 2, 4, 6, 8, 12, 16, or 24 hours (tick marks; pIC, poly(I:C); GRD, gardiquimod). Replicates were collapsed and genes hierarchically clustered (rows, genes; columns, experiments; red, induced from baseline; blue, repressed from baseline; white, unchanged from baseline).

(B) Model illustrating the differential gene regulatory networks controlling the antiviral ("poly(I:C)-like") and inflammatory ("PAM-like") programs.

The observed transcriptional responses were classified into a "PAM-like"' program and a "poly(I:C)-like" program, as well as a shared response (24.5% shared by PAM/poly(I:C)/LPS). The LPS response was largely the union of the "PAM-like" and "poly(I:C)-like" programs. This is partly explained by the known signaling pathways activated by these agonists. PAM binds TLR2 and signals through the MYD88 pathway; poly(I:C) binds TLR3 and MDA-5 and signals mostly through the TRIF and IPS-1 pathways, respectively; and LPS binds TLR4 and co-receptors and uses the MYD88 and TRIF pathways (Kagan et al., 2008). It is also consistent with the known induction of an anti-viral response by poly(I:C) and LPS (Doyle et al., 2002). The "PAM-like" program is enriched for the transcription factor NF-κB and for inflammatory responsive genes ($P < 6.1 \times 10^{-8}$), whereas the "poly(I:C)-like" program is enriched for interferon regulatory factors (IRFs) and for viral- and interferon-responsive genes (ISGs, $P < 8.3 \times 10^{-24}$). For simplicity, we thus refer to them as the "inflammatory" and "antiviral" programs (**Figure 2.2B**). A small number of genes are specific to a single stimulus. For example, ~250 genes are poly(I:C)-specific (1250 are shared with LPS), including several Type I interferons (*e.g.*, *Ifna2*, *Ifna4*, **Figure 2.2A**). Surprisingly, 82% of the gardiquimod (TLR7) and CpG (TLR9) response was shared with the LPS response, but with a weaker anti-viral component (**Figure 2.2**). This observation is unexpected given their different signaling mechanisms (Kawai et al., 2004), but is highly reproducible and robust (Amit et al., 2009).

2.5. Identification of Candidate Regulators and a Response Signature

To select potential regulators that mediate the observed transcriptional response, we focused on regulator genes whose expression changes during pathogen sensing, which is a reasonable assumption for many mammalian responses (Amit et al., 2007; Pe'er et al., 2002), including pathogen-sensing (Huang et al., 2001; Ramsey et al., 2008). First, we reconstructed an observational *trans*-model of gene regulation that associated 80 modules of co-regulated genes

with 608 predictive regulators (Lee et al., 2009; Ramsey et al., 2008; Zou and Hastie, 2005), automatically chosen out of a curated list of 3287 candidate regulators. This draft observational model was generated from the transcriptional profiles using a modification of the Module Networks algorithm (Zou and Hastie, 2005). Briefly, we used a probabilistic graphical model that takes an expression dataset and a set of candidate regulators (*e.g.*, all known and putative transcription factors in the genome), and reconstructs modules of co-regulated genes, and a regulation program that attempts to explain the expression of each module as a function of the expression of specific regulators. Here, we used a curated list of 3287 candidate proximal regulators of mRNA levels, including 1885 transcription factors, 1069 RNA binding proteins and 333 chromatin factors. Filtering identified 117 regulators above a minimal expression signal in at least one experiment. These included known regulators from the NF-κB, STAT, and IRF families, as well as unexpected candidates such as the circadian regulator Timeless and the DNA methyltransferase Dnmt3a. Second, we added five constitutively expressed regulators (Irf3, Rela, Nfe2l2, Ets1, Creb3) whose *cis*-regulatory elements are enriched in the responsive genes. Third, to capture delayed responses or nonlinear relations, we incorporated 22 regulators whose expression changed by at least a factor of 2. This resulted in 144 candidate regulators, with a distribution of expression patterns similar to the general response (Amit et al., 2009). The regulators' expression under LPS was conserved between DCs and functionally similar macrophages (Pearson correlation r = ~0.9 at 1 h) as well as between human macrophages and mouse DCs (r = ~ 0.6 at 2 h), supporting the functional relevance of the regulators' transcription (Amit et al., 2009).

To identify highly informative reporter genes for monitoring the effects of perturbing regulators, we devised an algorithm referred to as *GeneSelector* (Amit et al., 2009). This algorithm incrementally chooses genes (from our full expression dataset) whose expression profile improves our discrimination of stimuli given the previously chosen genes. Using this approach,

we identified the optimal time point (six hours post activation), and a set of 81 genes that distinguishes the stimuli. We added 37 candidate regulators with detectable expression at the 6-hour time point, creating a signature set of 118 genes. Finally we added 10 control genes whose expression levels were unchanged under all stimuli, but whose basal, constant levels varied from very low to high.

2.6. Perturbations, Profiling and Modeling

We generated validated lentiviral shRNAs that knocked down expression of 125 of our 144 candidate regulators by at least 75% and 32 shRNAs with no known gene targets as controls in bone marrow DCs (Amit et al., 2009). To carry out our perturbational study, we selected a single treatment, LPS, that activates the majority of both the inflammatory and antiviral programs. After stimulation of shRNA-perturbed DCs with LPS for 6 hours, we used a multiplex mRNA counting methodology to quantify the transcript levels of the 118 reporter and 10 control genes (Geiss et al., 2008).

The changes in signature gene expression resulting from infection with each shRNA were used to construct a model that associated regulators to their targets. We expected increases in the transcript levels of reporter genes whose repressors are targeted by knockdown, and decreases in reporters whose activators are targeted. Our false discovery rate (FDR) model estimates the statistical significance of a change in transcripts in DCs infected with a given shRNA (**Experimental Procedures**). We controlled for gene-specific noise by comparing to changes in the expression of each gene after perturbation with the control shRNAs (**Figure 2.3A**), and for shRNA-specific noise by comparing to changes in the expression of the control genes after a given shRNA perturbation (**Figure 2.3B**). We estimated the sensitivity of our calls from the 37 regulators, which are also included as target reporters (Amit et al., 2009).

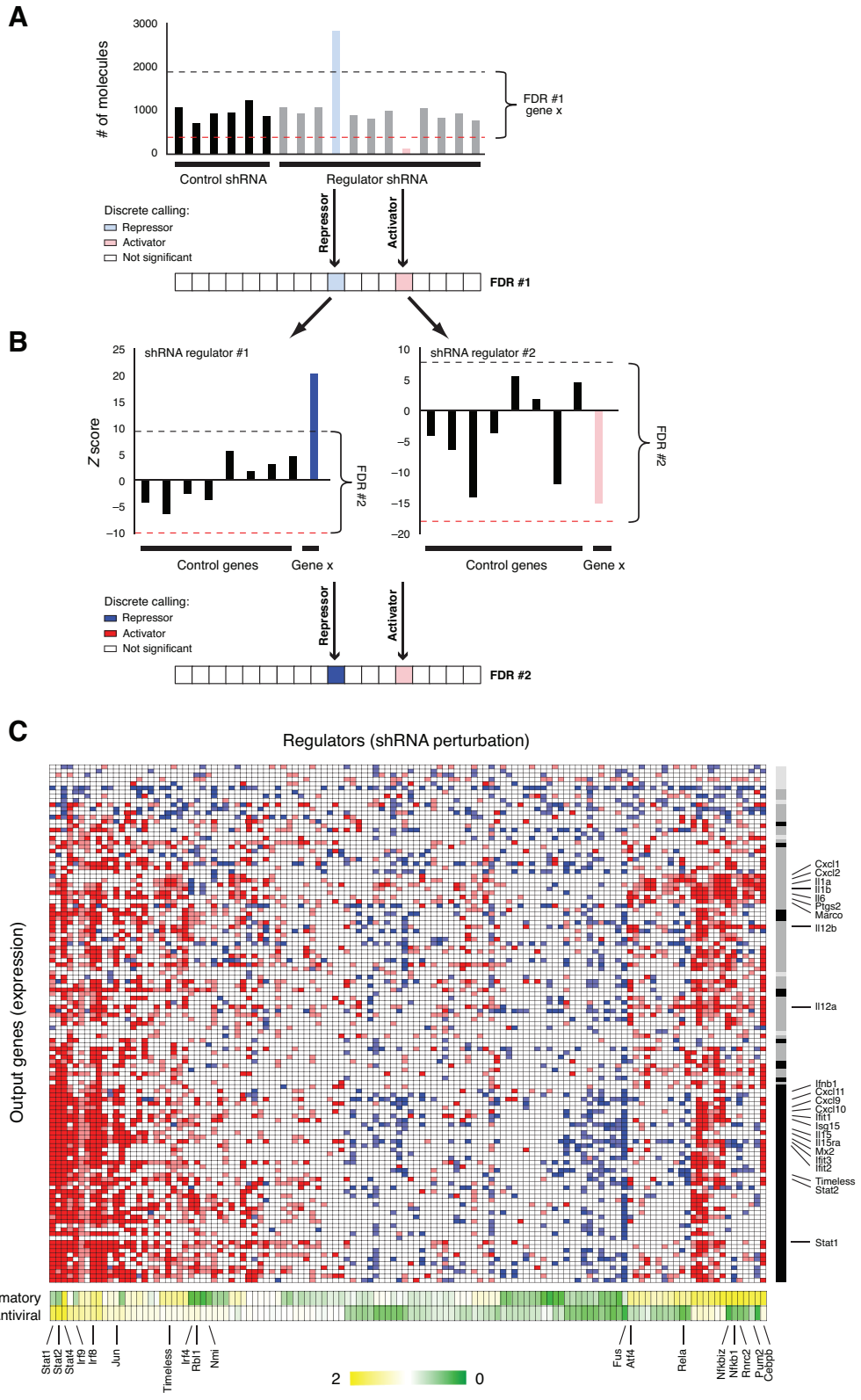

Figure 2.3. Gene Regulatory Programs Controlling the Response to Pathogen Components

(A-B) A strategy to minimize the false discovery rate (FDR) calls of significant changes in an output target gene resulting from knockdown of a regulator gene. (A) The first FDR procedure (top) compares the

Figure 2.3 (Continued)
expression of the gene after a perturbation with a regulator shRNA (right) to its expression upon perturbation with 32 nontargeting shRNAs (left). The dashed lines identify the gene-specific FDR-based thresholds for induction (blue line) and repression (red line). A discrete vector of significant calls (bottom) is derived from the raw data (blue, regulator represses the target gene; red, regulator induces the target gene). (B) A second FDR procedure (top) compares the expression of the target gene to that of eight control (target) genes upon perturbation with the same shRNA. In the example shown, the gene's induction (left) was significant relative to the variation in expression among the control target genes, resulting in a high score (bottom, dark blue), but its repression did not significantly differ from the controls, resulting in a lower score (bottom, weaker red).
(C) All significant (95% confidence) relations between regulators (columns) and targets (rows), colored as in (B). The gray bar (right) represents the NMF-based calls for each target gene; black, antiviral program; dark gray, inflammatory program; light gray, control genes. The bottom bar shows the degree of effect by the regulator on each program as determined by the NMF projection of the regulator's perturbation profile (yellow, high; green, low).

On the basis of these results, we identified a densely overlapping network with 2322 significant regulatory connections, including 1728 activations and 594 repressions (**Figure 2.3C**, red and blue, respectively, at 95% confidence; Amit et al., 2009). Of the 125 tested regulators, we confidently identified 100 with at least four targets. Among those were 24 hub regulators that were predicted to regulate more than 25% of the 118 genes measured, as well as 76 specific regulators, each affecting the expression of 4 to 25 genes. On average, ~14 (± 8 SD) regulators activated a target gene, and 5 (± 5.8 SD) regulators repressed it. Indirect effects may account for the large number of regulators we observed for each target.

Our perturbational model captured known regulatory features of the response and identified novel regulators. The reporter genes partition into two main clusters based on their response to perturbations (**Figure 2.3C**), consistent with the expression data: the antiviral ("poly(I:C)-like") program reporters (*e.g., Cxcl10, Isg15, Ifit1*), and the inflammatory ("PAM-like") program reporters (*e.g., Il1b, Cxcl2, Il6, Il12b*). We also found many known regulatory relations – for example, the NF-κB family of transcription factors (Rel, Rela, Relb, Nfkb1, and Nfkb2) regulating their known inflammatory gene targets. Our network provided evidence for the involvement of at least 68 additional regulators in the response to pathogens, of which 11 were hubs not previously associated with this system. Interestingly, 12 regulators identified (*e.g.,* Hhex, Fus,

Bat5, Pa2g4) are in linkage disequilibrium with single nucleotide polymorphisms (SNPs) associated with autoimmune and related diseases in genome-wide association studies (Amit et al., 2009).

2.7. The Core Inflammatory and Antiviral Programs

We next addressed how each regulator contributes to the generation of specific cell states. We first automatically defined the two major states induced by the five pathogen components with the use of non-negative matrix factorization (NMF) (Brunet et al., 2004) and the original array data (Amit et al., 2009). This procedure identified two major expression components (termed "metagenes"): one predominantly determined by genes from the inflammatory program and the other by genes from the antiviral programs (**Figure 2.2A**). Next, we quantified the effects of each regulator's knockdown on these two states (**Figure 2.3C**, Amit et al., 2009) by classifying the target mRNA expression measurements following a regulator's perturbation (Brunet et al., 2004; Daily et al., 2007).

Finally, we used a regulator ranking score (Amit et al., 2009) to assign 33 (including 8 known) genes as regulators of the inflammatory state and 33 (including 15 known) genes as regulators of the antiviral state. This accurately classified the known activators of the inflammatory response (*e.g.*, NF-κB and related factors Rela, Nfkbiz, Nfkb1, **Figure 2.3C**, yellow in the inflammatory metagene), and of the antiviral response (*e.g.,* Stat1, Stat2, Stat4, Irf8, Irf9; **Figure 2.3C**, yellow in the viral metagene). Although all perturbation experiments were conducted only under LPS stimulation (a bacterial component), we correctly classified factors known to mediate the response to other stimuli. Because 34 additional regulators were associated with both responses, it is possible that a single regulator can control genes in either state, depending on the differential timing of regulator activation, its level, or combinatorial regulation. Notably, for 12

of the transcription factors examined, we found an enriched *cis*-regulatory element in the appropriate metagene (Amit et al., 2009).

On the basis of the NMF scores, we identified an inflammatory subnetwork (Amit et al., 2009), an antiviral subnetwork (**Figure 2.4A**), and several fine-tuning subnetworks that affect smaller numbers of genes from both responses (Amit et al., 2009). The inflammatory subnetwork consisted of three regulatory modes: dominant activators (Cebpb, Bcl3, Cited2) which induce more inflammatory targets than anti-viral ones; cross-inhibitors (Nfkbiz, Nfkb1, Atf4, Pnrc2) which induce inflammatory genes while repressing anti-viral ones, and specific activators (Runx1, Plagl2), that only target inflammatory genes (Amit et al., 2009). We observed that dominant activators mostly regulate effectors, whereas regulators are primarily controlled by cross-inhibitors.

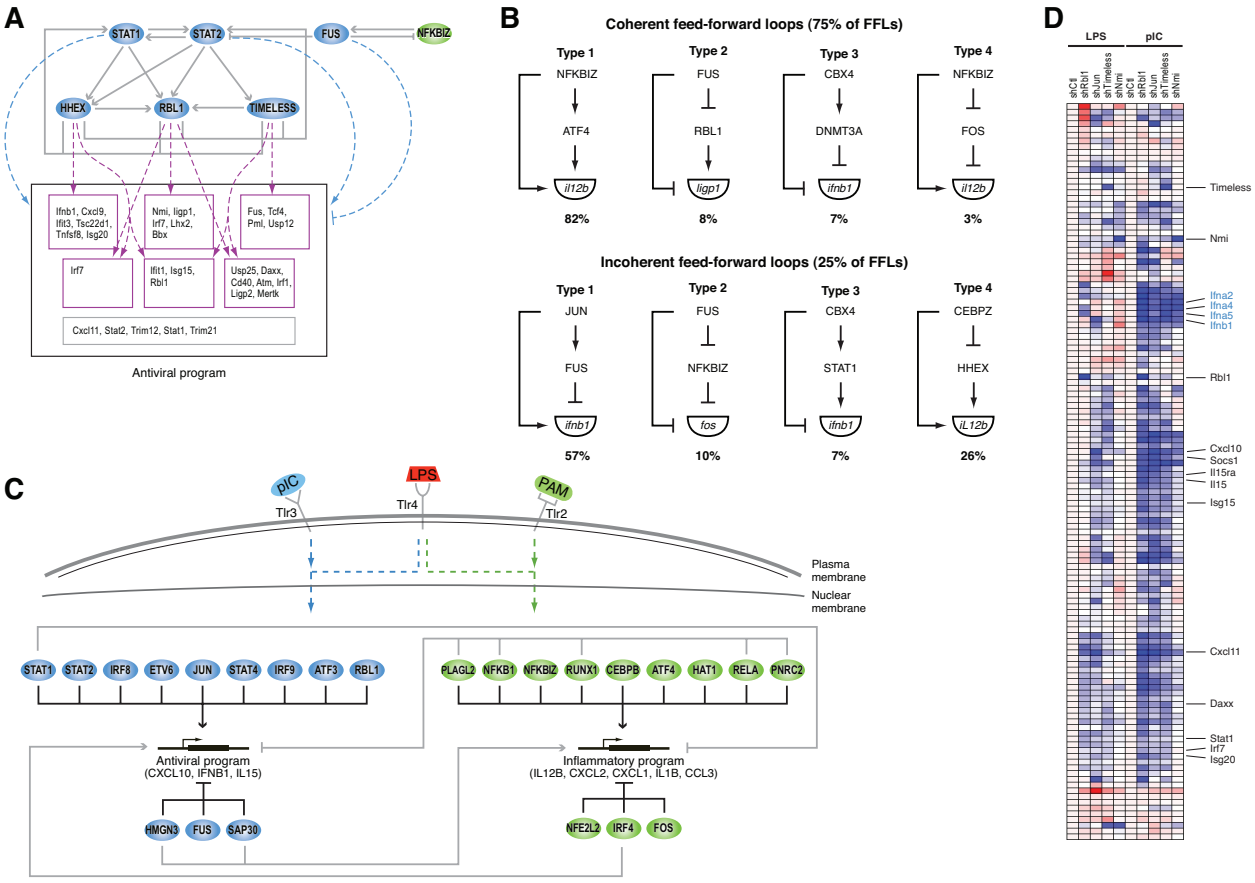

Figure 2.4. The Core Regulatory Circuits Controlling the Inflammatory and Antiviral Responses

(A) The antiviral subnetwork shows regulatory relations between the core antiviral regulators (blue nodes, top), their targets (boxes, bottom), each other, and inflammatory regulators (green node, top right). The two top regulators, Stat1 and Stat2, activate all antiviral targets (dashed blue arrows). The second-tier regulators activate subsets of targets (dashed purple arrows).

(B) Examples of feedforward loop classes identified in the network, with fraction of each class.

(C) A core regulatory network of the inflammatory and antiviral programs, consisting of the most distinct regulators, and their relation to ligands and receptors (top). Pointed arrows, induction; blunt arrows, repression; green ovals, inflammatory regulators; blue ovals, antiviral regulators. Example target genes are noted.

(D) mRNA expression profiles for the TLR target genes (rows) upon perturbation with shRNAs against a subset of viral regulators (columns) and followed by stimulation with LPS (left) or poly(I:C) (right). All values are normalized by expression in cells infected with a control shRNA and under the same stimulus (shCtl).

Focusing on the network architecture, we found multiple feedforward circuits in this response, where an upstream regulator controls a target gene both directly and indirectly through a secondary regulator (**Figure 2.4B**) (Amit et al., 2009; Mangan and Alon, 2003). The majority (~75%, 4892 of 6444) of these feedforward circuits were found to be coherent (Mangan and Alon, 2003), having the same direct and indirect effect on the regulated gene. The vast majority

(80%) are type I loops (Alon, 2007) with all-positive regulation (*e.g.,* Nfkbiz activates E2f5 and both activate IL6). Such feedforward circuits respond to persistent rather than transient stimulation, protecting the system from responding to spurious signals, as was shown for one circuit in LPS-stimulated macrophages (Litvak et al., 2009). Our finding suggests that coherent feedforward loops, especially class I (Mangan and Alon, 2003), are a general design principle in this system and may physiologically impact this response.

In the antiviral sub-network, we identified a two-tiered regulatory circuit combining feed-forward and feedback loops (**Figure 2.4A**) (Amit et al., 2009). This circuit has at the top the antiviral regulators Stat1 and Stat2, which regulate a full complement of anti-viral reporters. The second-tier regulators, Timeless, Rbl1 and Hhex are controlled by Stat1 and Stat2 and most likely form coherent feedforward loops that target specific subsets of genes. Timeless, Rbl1 and Hhex also feed back and promote the expression of the Stat regulators. This circuit is repressed through the cell cycle regulator and RNA binding protein Fus (Wang et al., 2008), acting as a single dominant inhibitor of 43 viral genes.

Finally, we derived a core network incorporating the regulators with the most substantial impact on each response, on the basis of the number, magnitude, and logic of targets that each regulator affects (**Experimental Procedures**). The core network (**Figure 2.4C**) has 24 regulators, 13 of which have previously been identified as key factors regulating the inflammatory or antiviral responses; the other 11 have not been previously implicated in either response. Of these, 19 are transcription factors, three are chromatin modifiers, and two are RNA binding proteins. The regulators apparently distinguish the two programs through cross-inhibition (**Figure 2.4C**, gray lines) or dominant activation (**Figure 2.4C**). The core network also explains how differential expression of secreted factors is specified, leading to the activation and migration of appropriate cell types for different pathogens (Amit et al., 2009; Luster, 2002).

Embedded within the many known regulators of the antiviral response (**Figure 2.4C**), we found a large set of regulators not previously associated with this response. These included several known regulators of the cell cycle and the circadian rhythm, including Rbl1, Jun, Rb, E2f5, E2f8, Nmi, Fus, and Timeless, several of which were placed in our core network. This suggests that a cell cycle regulatory circuit was co-opted to function in the antiviral response in DCs (with no observable effect on cell cycle progression; Amit et al., 2009). Since we identified these antiviral regulatory relations in perturbation experiments using DCs stimulated with the bacterial component LPS, we silenced four regulators (Timeless, Rbl1, Jun and Nmi) after exposure to the viral component poly(I:C). Each of the four regulators had a strong impact on the antiviral program, more than was observed under LPS stimulation (**Figure 2.4D**), and on affected genes (*e.g.,* Type I IFNs) whose expression is poly(I:C)-specific. Nmi affected a smaller set of genes, consistent with the model's prediction. These results demonstrate our ability to correctly predict function in unobserved conditions.

Although most antiviral genes are induced following stimulation with the bacterial component LPS, a few critical ones are expressed specifically in poly(I:C) stimulation, or follow distinct patterns in each stimulus. In response to viral infection cells induce the production of interferon β (*Ifnb1*), a crucial mediator of the antiviral response. Because high levels of interferon β may be deleterious to the host if infected by specific bacteria (Decker et al., 2005), we predicted that specific mechanisms insulate *Ifnb1* mRNA regulation from the response to LPS. Indeed, although *Ifnb1* expression was induced in the first two hours of stimulation with LPS, this expression declined at subsequent time points, in contrast to its sustained induction following poly(I:C) treatment (**Figure 2.5A**). Our model suggested that three regulators known to affect chromatin remodeling are *Ifnb1* repressors in LPS (**Figure 2.5B**): the Polycomb complex subunit Cbx4 (Bernstein et al., 2006), Fus (Wang et al., 2008), and the DNA methyltransferase

Dnmt3a (Li et al., 2007). Cbx4 appeared to confer antiviral specificity to *Ifnb1* mRNA induction as it is induced within the first two hours of PAM and LPS treatments but not by poly(I:C) (**Figure 2.5C**). Cbx4 knockdown caused induction of *Ifnb1* mRNA and protein during LPS treatment (**Figure 2.5D**), but had no effect on the induction of the chemokine *Cxcl10*, a poly(I:C)- and LPS-induced gene (Amit et al., 2009). Cbx4 knockdown did not affect *Ifnb1* during PAM activation (**Figure 2.5E**), when the antiviral response is not induced. Combined with evidence for chromatin changes around the *Ifnb1* locus and its closest neighbor gene, *Ptplad2*, which has a similar dependence on Cbx4, these data are consistent with an effect by Cbx4 on local chromatin organization (Amit et al., 2009). Cbx4 knockdown affected few genes (~120 up-regulated and ~120 down-regulated genome-wide) (Amit et al., 2009). Because most up-regulated genes show a precise temporal pattern in unperturbed cells akin to that of Cbx4 – they are induced quickly and return to basal level by 2-4 hours (Amit et al., 2009) – we conclude that a chromatin modifier can act like a transcription factor controlling the precise expression of specific genes in the regulatory program.

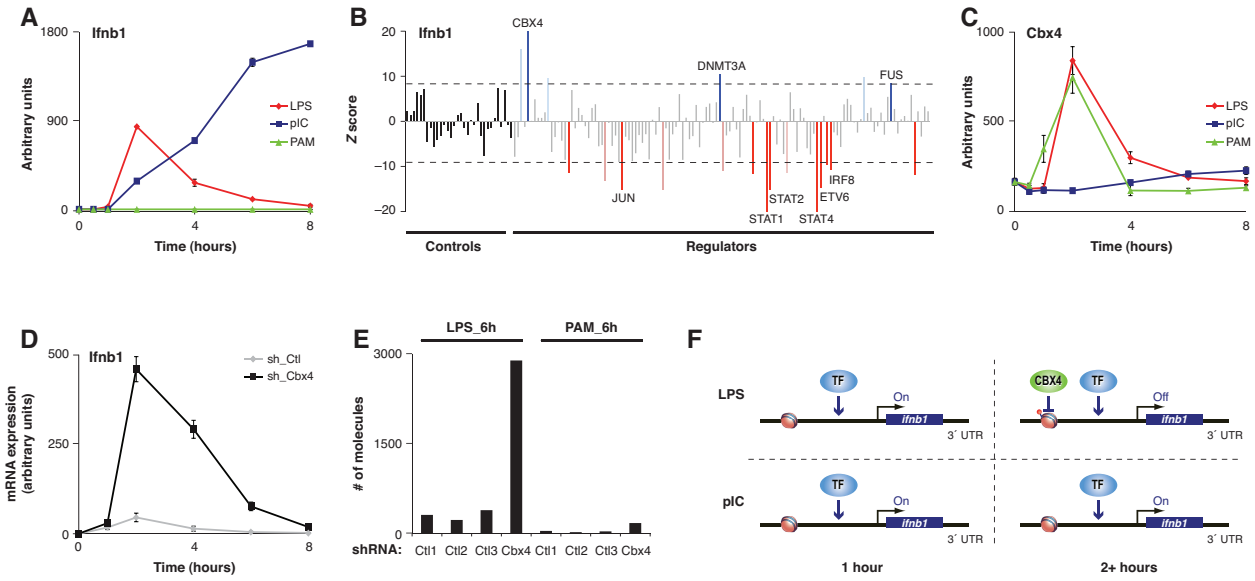

Figure 2.5. The Polycomb Component Cbx4 Selectively Restricts *Ifnb1* Production upon Stimulation with Bacterial Components

(A and C) LPS (red), poly(I:C) (blue), and PAM (green)–induced expression of *Ifnb1* (A) and *Cbx4* (C) derived from data in Fig. 2A.

(B) *Ifnb1* expression (by nCounter) in response to LPS in DCs perturbed by control shRNAs or shRNAs targeting each of 125 regulators (format as in Figure 2.3B).

(D) *Ifnb1* mRNA levels (by quantitative reverse transcription polymerase chain reaction) after LPS treatment in unsorted mouse bone marrow DCs perturbed with an shRNA against Cbx4 (black) or a control shRNA (gray); signals are relative to t = 0.

(E) *Ifnb1* mRNA levels (by nCounter) at 6 hours after exposure to LPS or PAM in bone marrow DCs perturbed with an shRNA against Cbx4 or one of three control shRNAs.

(F) Model for bacterial-specific repression of *Ifnb1* by Cbx4. Both poly(I:C) and LPS induce *Ifnb1* expression early (left), but only LPS induces *Cbx4*, which then represses the *Ifnb1* locus at a later time (right, top).

Taken together, our results suggest a model of a transcriptional negative feedback loop, controlling *Ifnb1* expression in LPS stimulation, wherein the induced pro-inflammatory regulator and chromatin modifier Cbx4 represses transcription by modifying the chromatin in the *Ifnb1* locus, generating the specificity needed to drive inflammatory versus the antiviral response (**Figure 2.5F**). The Type III coherent feedforward loop formed by Cbx4 and Dnmt3a (**Figure 2.4B**) is consistent with a delayed repression of *Ifnb1*. Because neither regulator carries a sequence-specific DNA binding domain, the factors responsible for their guidance to the *Ifnb1* locus remain unknown.

2.8. Discussion

A central goal of our study was to address the mechanistic basis for pathogen-specific responses. Consistent with previous studies (Doyle et al., 2002), we distinguished two key programs, a “PAM-like” inflammatory response and a “poly(I:C)-like” antiviral response, which are together induced by LPS, a Gram-negative bacterial component recognized by TLR4. These programs reflect both qualitative and quantitative differences between the required functional responses, and are consistent with the cross-protection between certain bacteria and virus infections (Doyle et al., 2002). The broad effect of LPS allowed us to focus on a single stimulus and time point, but screens with other stimuli may identify additional unique regulators.

We found these two responses to be controlled by two corresponding regulatory arms, uncovering a mechanistic basis for the observed transcriptional responses. These two arms are integrated into a core network of two dozens regulators which balances specific and shared responses through dominant activation and cross-inhibition. In the inflammatory response, we found several feedforward loops, which may ensure response to only persistent, but not sporadic, signals. In the antiviral response, we discovered a two-tiered circuit involving feedback and feedforward loops, implicating a module of cell cycle regulators (Jun, Rbl1, Timeless, and Nmi), which we directly validated. Over 75 additional genes work to further fine-tune the regulation of gene targets. This perturbational model identifies many regulatory relations that would have been missed by non-systematic approaches.

Our work establishes an unbiased, straightforward, and general framework for network reconstruction in mammalian cells (Amit et al., 2009). This approach can be executed at substantial scale and reasonable cost, and is compatible with the challenge of deciphering the multiple regulatory systems that operate in mammals. It can be expanded to derive increasingly detailed models, and distinguish direct from indirect targets. It will also facilitate the

development of new computational approaches to infer regulatory models. Although many computational approaches have attempted to derive observational models, their quality has been difficult to evaluate (Kim et al., 2009). The data generated here includes both expression profiles for training a model, as well as a perturbational unbiased screen for testing its quality (ftp://ftp.broadinstitute.org/pub/papers/dc_network/). When we compared the perturbational model to our observational model, we found that many candidate regulators were correctly identified in both; however, there were also numerous false positive relations in the observational model, attributable to the fact that both the correct regulator and many others have indistinguishable expression (Amit et al., 2009).

The high-resolution map we constructed might have biomedical implications. By identifying regulators that mediate the differential control of specific gene pairs (*e.g.*, IL-23 vs. IL-12) and entire regulatory arms (*e.g.,* viral vs. inflammatory), it opens the way for therapeutic targeting of specific pathways to control disease or enhance vaccine efficacy. Furthermore, 12 of our regulators reside in genetic loci that were in linkage disequilibrium with SNPs associated with autoimmune and related diseases. The identified genes and their impact on DCs provide hypotheses to help explain how alleles of genes in a cascade may alter susceptibility to specific infections or immune disorders in humans.

2.9. Experimental Procedures

Preparation of dendritic cells

Bone marrow-derived dendritic cells (BMDCs) were generated from 6-8 week old female C57BL/6J mice (Jackson Laboratories). Bone marrow cells were collected from femora and tibiae and plated at 10^6 cells/mL on non-tissue culture treated petri dishes in RPMI-1640 medium (Gibco), supplemented with 10% FBS, L-glutamine, penicillin/streptomycin, MEM non-

essential amino acids, HEPES, sodium pyruvate, β-mercaptoethanol, and murine GM-CSF (15 ng/mL; Peprotech) or human Flt3L (100 ng/mL; Peprotech). GM-CSF-derived BMDCs were used directly for all RNAi experiments unsing lentiviral shRNAs. For all other experiments, floating cells from GM-CSF cultures were sorted at day 5 by MACS using the CD11c (N418) MicroBeads kit (Miltenyi Biotec). Sorted $CD11c^+$ cells were used as GM-CSF-derived BMDCs, and plated at 10^6 cells/mL and stimulated at 16 h post sorting.

TLR agonists

TLR ligands were from Invivogen (Pam3CSK4, ultra-pure *E. coli* K12 LPS, CpG) and Enzo Life Sciences (poly(I:C)), and were used at the following concentrations: Pam3CSK4 (250 ng/mL), poly(I:C) (10 μg/mL), LPS (100 ng/mL), gardiquimod (250 ng/mL), CpG (1 μg/mL).

mRNA isolation

Total RNA was extracted with QIAzol reagent following the miRNeasy kit's procedure (Qiagen), and sample quality was tested on a 2100 Bioanalyzer (Agilent). RNA was reverse transcribed with the High Capacity cDNA Reverse Transcription kit (Applied Biosystems). For experiments with more than 12 samples, we harvested polyA+ RNA in 96- or 384-well plates with the Turbocapture mRNA kit (Qiagen) and reverse transcribed with the Sensiscript RT kit (Qiagen).

qPCR measurements

Real time quantitative PCR reactions were performed on the LightCycler 480 system (Roche) with FastStart Universal SYBR Green Master Mix (Roche). Every reaction was run in triplicate and GAPDH levels were used as an endogenous control for normalization.

Microarray hybridization and processing

For oligonucleotide microarray hybridization, 1.5 μg RNA were labeled, fragmented and hybridized to an Affymetrix Mouse Genome 430A 2.0 Array. After scanning, the expression

value for each gene was calculated with RMA (Robust Multi-Array) normalization (Irizarry et al., 2003). The average intensity difference values were normalized across the sample set. Probe sets that were absent in all samples according to Affymetrix flags were removed. All values lower than 50 were floored to 50. Only probe sets that changed in at least one pair of biological duplicates by 1.7 fold or more were analyzed further in this study. We defined induced probesets for each condition (TLR agonist) as probesets that display at least 1.7 fold up-regulation in both duplicates of at least one time point, as compared to the control. Control values were defined as the mean expression of the control sample, calculated over control non stimulated samples times 0, 1, 2 and 4 hours.

Regulatory network reconstruction

We reconstruct a draft model with a modification of the Module Networks algorithm. We first applied Module Networks as original developed (Segal et al., 2003). Briefly, we use an iterative learning procedure using the Expectation Maximization (EM) algorithm. Each iteration consists of two steps: an E-step and an M-step. In the M-step, the procedure is given a partition of the genes into modules and learns the best regulation program (as a regression tree) for each module (for efficiency, some M-steps only optimize the parameters of the normal distributions at the leaves of the regulation tree). In the E-step, given the inferred regulation programs, we re-assign each gene to the module that best predicts the gene's behavior (we do not assign a regulator gene to a module in which it is also a regulatory input, directly or indirectly). The regulation program is learned from a pre-defined set of candidate regulators (3287 TFs, CFs and RNA-BPs). We applied Module Networks systematically when initialized to 10-250 modules (in increments of 10), and chose the model whose likelihood score was 70% of the best score. The chosen model consisted of 80 modules.

The regression tree in the original model suffers from two potential limitations: it allows only a single regulator in each split in the tree (even when several equally good ones exist), and it does not allow for time lags. To address these limitations, we next refined the regulatory programs of the 80 modules using the Elastic Net (Zou and Hastie, 2005), an L2-regularized linear regression procedure. For each module, we used the LARS-EN algorithm (Zou and Hastie, 2005) to regress the mean profile of the module's genes with a selected combination of candidate regulators (to avoid cyclicity, we eliminated the regulator genes from the modules in this step). In addition, we allowed up to a fourth time point "lag" between the expression of the regulators and that of the target module. Finally, we used a bootstrap procedure, holding off 20% of each module's genes at a time and repeating the learning (Amit et al., 2009).

Custom Nanostring CodeSet construction using the GeneSelector algorithm

To choose a set of genes that will capture as much as possible of the information on the expression of all genes, we used an information-theoretic approach. We modeled the expression levels *X* given the experimental condition *C* with a naïve Bayes model where the expression of gene *i* under condition *c* follows a normal distribution $X_i \mid C = c \sim \mathrm{N}(\mu_{ic}, \sigma_i^2)$. In this model, the expression levels of all genes depend on the experimental condition *C*, and we selected genes that are highly informative about *C*. Formally, for a set of genes *Y* we used the conditional entropy $H(C \mid \mathbf{Y}) = -\sum_c p(C=c) \sum_y p(Y=y \mid C=c) \log p(C=c \mid Y=y)$ as a measure of the remaining uncertainty in *C* once the expression levels *Y* are known. We then used this measure and a greedy procedure to select multiple disjoint gene sets, $Y_1, \ldots, Y_k$ such that for each set *Yi*, $H(C \mid Y_i) < \eta$ (we set η = 0.5). In the greedy procedure, we select genes one at a time, and with each selected gene re-compute the entropy given the genes already selected in the current set. Once a set is complete (the remaining conditional entropy is below the threshold η), we add all the genes to the selected set, and repeat the procedure (excluding

all the selected genes from consideration). We stop when the number of selected genes has reached a user-defined threshold, set by the number of genes feasible for the experimental assay.

To select a time point, we used the same approach. Here, we measured entropy under all time points for multiple randomly selected gene sets of several sizes and plotted the average entropy for each timepoint (Amit et al., 2009). We chose the time point with the minimal entropy (*i.e.*, 6 h after LPS stimulation).

Analysis of cis-regulatory elements

Cis-regulatory elements were represented by a Position Weight Matrix (PWM). We compiled a set of 1651 PWMs from the TRANSFAC matrix database v8.3 (Matys et al., 2006), JASPAR Version 2008 (Sandelin et al., 2004) and experimentally determined PWMs (Badis et al., 2009; Berger et al., 2008). Given a PWM, for each nucleotide position in the promoter of each mouse gene, we calculated an affinity score defined as the log-likelihood ratio (LOD score) for observing the sequence given the PWM versus a given random genomic background. We then found the best conserved motif instance over the entire promoter region. We automatically computed a PWM-specific cutoff, by using the information content of each motif, computed as the 2-IC quantile of the PWM LODs distribution. We considered a “hit” in the promoter if the maximal LOD score was above this cutoff. Finally, we computed the enrichment of the motif in each of six clusters determined by the microarray experiments, using a two-sided Wilcoxon rank-sum test between the set and the background (Amit et al., 2009). To ensure that enrichment was not due to nucleotide bias within the promoter, we also shuffled the PWM and computed enrichment for the true PWM compared to the shuffled PWMs. A motif was

considered enriched in a gene set if it passed *p*-value < 0.01. We included a PWM in our final set of regulators if it was associated with at least one gene set.

shRNA experiments

High titer lentiviruses encoding shRNA-targeting genes of interest were obtained from The RNAi Consortium (TRC; Broad Institute, Cambridge, MA, USA) (Moffat et al., 2006). 10^5 bone marrow cells were plated in 100 µL complete RPMI on non-tissue culture treated flat bottom 96-well plates (Nunc). At day 2, cells were spin-infected (45 min, 37°C, 2200 rpm) with 20 µL of shRNA-encoding lentiviruses in the presence of polybrene (8 µg/mL), and subsequently fed with 100 µL complete RPMI with GM-CSF. At day 4, infected cells were selected by adding puromycin (5 µg/mL) to the culture in 25 µL complete RPMI with GM-CSF. Infected cells were used for analysis 3 days after initiating puromycin selection. For each regulator, we tested five shRNAs for knock down efficiency using qRT-PCR of the target gene. Lentivirus-infected cells were composed of ~90% DCs (Amit et al., 2009), which was highly similar to sorted $CD11c^+$ DCs.

mRNA measurements with nCounter

Details on the nCounter system are presented in full in (Geiss et al., 2008). CodeSets were constructed to detect genes selected by the GeneSelector algorithm and additional controls as described. ~5 × 10^4 bone marrow cells were lysed in RLT buffer (Qiagen) supplemented with 1% β-mercaptoethanol. 10% of the lysate was hybridized for 16 hours with the codeset and loaded into the nCounter prep station followed by quantification using the nCounter Digital Analyzer.

Normalization of nCounter data

We normalized the nCounter data in three steps. In the first step, we controlled for small variations in the efficiency of the automated sample processing. To this end, we followed the

manufacturer's instructions, and normalized measurements from all samples analyzed on a given run to the levels of a chosen sample (in all cases we used the first sample in the set). This was done using the positive spiked-in controls provided by the nCounter instrument. In the next step, we relied on ten control genes (*Ik, Ndufa7, Tomm7, Tbca, Ndufs5, Ywhaz, Mea1, Rbm6, Shfm1, and Gapdh*), which were included as reporters and were identified from the microarray experiment as unchanged upon stimulation by any of the pathogen components. We found that two of these genes (*Gapdh* and *Rbm6*) showed too much variation and removed them from all subsequent analysis. We used the remaining eight genes for a second round of normalization. For every sample, we computed the weighted average m_i of the mRNA counts of the seven transcripts and normalized the sample's values by multiplying by the constant m_1/m_i. Finally, we obtained a normalized expression quantity that takes into account the intrinsic noise in our system. We used the 32 samples treated with control shRNAs (that do not target any gene in the mouse genome) to define a *z*-statistic *z* for each observation o_{ij} of transcript *i* in each shRNA experiment *j*: $z_{ij} = \frac{o_{ij} - m_j}{s_j}$ where m_j and s_j are, respectively, the mean and variance of the expression of transcript *j* in the control shRNA experiments.

Confidence estimates for differential expression in perturbation experiments

We used two permutation-based approaches to estimate our confidence in an observed *z*-score value for a transcript in an shRNA experiment. In the first approach, we defined a per-gene confidence score for each measurement, by using the variation in that gene's expression in the control shRNA experiments. We computed the confidence scores for each measurement (one gene in one experiment) at a time, by swapping the measured value of that gene with each of its measurements in the control experiments in turn, and re-calculating a new *z*-score. We then assessed the significance of the real *z*-score, given the distribution of the permuted scores as a null distribution. More formally, for each of the observed counts o_{ij} of the reporter gene *i* in

sample *j* we generated *r* permuted values (where *r* is the number of control shRNA experiments) as follows. Let $c_{i1}, \ldots, c_{ir}$ be the *r* transcript counts for gene *i* in each of the *r* control experiments. The *k* permuted *z*-score is obtained by swapping o_{ij} with c_{ik} and computing a *z*-score as $z_k^{ij} = \frac{c_{ik} - m_{ij}^k}{s_{ij}^k}$, where $m_{ij}^k = \frac{c_{i1} + \cdots + c_{ik-1} + o_{ij} + c_{ik+1} + \cdots + c_{ir}}{r}$ and $s_{ij}^k = \frac{\sqrt{(c_{i1} - m_{ij}^k)^2 + \cdots + (c_{ik-1} - m_{ij}^k)^2 + (o_{ij} - m_{ij}^k)^2 + (c_{ik+1} - m_{ij}^k)^2 \cdots + (c_{ir} - m_{ij}^k)^2}}{\sqrt{r}}$. We take the permuted scores z_k^{ij} as a null distribution and the FDR for a given *z*-score z_{ij} for gene *i* in experiment *j* is given as $FDR(z) = \frac{E_k(\#\{z_k^{ij} \mid z_k^{ij} > z; j = 1,\ldots,n\})}{\#\{z_{ij} > z; j = 1,\ldots,n\}}$, where *n* is the number of shRNA experiments. The confidence for z is $conf(z) = 1 - FDR(z)$.

In the second approach, we devised a per-experiment confidence score for each measurement. We use a similar procedure to control the FDR on the *z*-statistic, based on variation in the expression of control genes in each experiment. Formally, let $z_{ij}, \ldots, z_{nj}$ be the *z*-scores for the j^{th} experiment (shRNA), and assume the first *l* transcripts are control transcripts whose expression does not change in response to any pathogen component (*l* = 8, see above). We defined $\tilde{z}_{ij} = \frac{z_{ij} - \tilde{m}_j}{\tilde{s}_j}$ where now $\tilde{m}_j$ and $\tilde{s}_j$ are, respectively, the mean and variance of the *z*-scores of the control transcripts 1, … , *l* in the j^{th} shRNA experiment. We perform *l* permutations as described above, by swapping each observed *z*-scores with a control transcript score and computing $\tilde{z}$, then computing an FDR as above.

NMF and scoring of regulator profiles

We performed NMF as previously described (Brunet et al., 2004). Briefly, we represented the microarray expression dataset as an $n \times N$ matrix, M, whose rows contain the expression levels of the n genes in the N samples. We first used NMF to we find $k=2$ metagenes, as we focused on two key responses. Each metagene is a positive linear combination of the N genes. To do this we find an approximate factoring, $M \approx W \times H$, where both factors have only positive entries. W is an $n \times k$ matrix that defines the metagene decomposition model. Its columns specify how much each of the n genes contributes to each of the k metagenes. We assign a target gene to be inflammatory or anti-viral based on the larger of its two scores in W. H is a $k \times N$ matrix whose entries represent the expression levels of the k metagenes for each of the N samples. From the factoring of M, we next constructed a mapping that allows us to project the perturbation profiles into the space of the metagenes. To do this we used the Moore–Penrose generalized pseudoinverse (Brunet et al., 2004) of W, such that, $\hat{H} = (W)^{-1} \times M$, where $\hat{H} \approx H$. We then applied the resulting pseudoinverse to each of the perturbation experiments to obtain a score for the experiment (and hence the perturbed regulators) on each of the two metagenes (Amit et al., 2009).

Regulator ranking score

Interactions between regulators and target genes that passed the two FDR thresholds (above) were selected. For each regulator, we counted the number of target genes, and separated the interactions on the basis of their sign (activation or repression) and their NMF metagene classification [F1 (inflammatory) or F2 (viral)]. We subtracted the number of repressed genes from the number of activated genes for each metagene, and scaled the numbers on the basis of the metagene 'size' (since the number of genes assigned to F1 and F2 is different). Regulators were ranked by the difference between the resulting two numbers (metagene F2 and metagene

F1) for each regulator. The cutoff threshold was a score greater then 2 or -2 (and >5 interactions).

Analysis of loci from genome wide association studies (GWAS)

We downloaded a list of disease-associated SNPs ($P < 0.05$) from the National Human Genome Research Institute catalog of published genome wide association studies (Hindorff et al., 2009). We then extracted regions which are in linkage disequilibrium (LD) with each of the SNPs by first finding the left-most and right-most SNPs that are in LD ($R^2 > 0.5$) with the aforementioned SNP and then finding the closet recombination hotspots as described in Raychaudhuri et al., 2009. We mapped the list of 125 regulators from the mouse genome (mm8; Mouse genome assembly 8, Feb 2006) to their orthologs in the human genome (hg17, Human genome assembly 17, May 2004) and identified those that resided within the aforementioned regions.

ChIP-sequencing

Dendritic cells were fixed with 1% formaldehyde, washed with ice-cold PBS, harvested by scraping, pelleted, and resuspended in SDS lysis buffer. Samples were sonicated with a Branson 250 Sonifier for 8 × 25 s cycles at 70% duty, output 3.5, centrifuged at 13,000 × g for 10 min, and diluted 10-fold in ChIP dilution buffer. After removing a control aliquot (whole-cell extract), the sample was incubated at 4°C overnight with antibodies against H3K4me3 (Abcam 8580) or H3K36me3 (Abcam 9050). Complexes were precipitated with protein A-Sepharose. Beads were washed sequentially with low-salt immune complex wash, high-salt immune complex wash, LiCl immune complex wash, and TE. Immunoprecipitated chromatin was eluted in elution buffer, incubated at 65°C for 8 h, and treated with proteinase K. DNA was purified by extracting twice with phenol and once with chloroform and precipitating in ethanol. The DNA was treated with RNase and purified with a MinElute Kit (Qiagen). Sequencing libraries were generated from 1–10ng of ChIP DNA by adaptor ligation, gel purification and 16 cycles of PCR.

Sequence reads from Illumina libraries were identified with standard Illumina basecalling software and then analyzed with a custom computational pipeline. Reads were aligned to the reference genome, and the fragment count at any given position (25-bp resolution) was estimated as the number of uniquely aligned reads oriented towards it and within 300bp.

ELISA

Cell culture supernatants were assayed using a sandwich ELISA kit for mouse IFN-β (PBL Biomedical Laboratories).

Fluorescence Activated Cell Sorting

Non-adherent cells were stained with A647 conjugated anti-CD11c (N418, Biolegend) and FITC labeled anti-I-A[b] MHC class II alloantigen (AF6-120.1, BD Pharmingen). Fc receptors were blocked with Mouse BD Fc Block (BD Pharmingen). Flow cytometry was done with a BD LSRII, and the data were analyzed with FlowJo software (Treestar).

Accession numbers and supplementary tables

Complete microarray data sets are available in the NCBI Gene Expression Omnibus under the accession number GSE17721. All 16 supplementary tables containing the datasets presented in this Chapter are available with the online version of this published work (Amit et al., 2009).

2.10. References

Alon, U. (2007). Network motifs: theory and experimental approaches. Nat Rev Genet *8*, 450-461.

Amit, I., Citri, A., Shay, T., Lu, Y., Katz, M., Zhang, F., Tarcic, G., Siwak, D., Lahad, J., Jacob-Hirsch, J.*, et al.* (2007). A module of negative feedback regulators defines growth factor signaling. Nat Genet *39*, 503-512.

Amit, I., Garber, M., Chevrier, N., Leite, A.P., Donner, Y., Eisenhaure, T., Guttman, M., Grenier, J.K., Li, W., Zuk, O., et al. (2009). Unbiased reconstruction of a mammalian transcriptional network mediating pathogen responses. Science *326*, 257-263.

Badis, G., Berger, M.F., Philippakis, A.A., Talukder, S., Gehrke, A.R., Jaeger, S.A., Chan, E.T., Metzler, G., Vedenko, A., Chen, X.*, et al.* (2009). Diversity and complexity in DNA recognition by transcription factors. Science *324*, 1720-1723.

Berger, M.F., Badis, G., Gehrke, A.R., Talukder, S., Philippakis, A.A., Pena-Castillo, L., Alleyne, T.M., Mnaimneh, S., Botvinnik, O.B., Chan, E.T.*, et al.* (2008). Variation in homeodomain DNA binding revealed by high-resolution analysis of sequence preferences. Cell *133*, 1266-1276.

Bernstein, E., Duncan, E.M., Masui, O., Gil, J., Heard, E., and Allis, C.D. (2006). Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. Mol Cell Biol *26*, 2560-2569.

Brunet, J.P., Tamayo, P., Golub, T.R., and Mesirov, J.P. (2004). Metagenes and molecular pattern discovery using matrix factorization. Proc Natl Acad Sci U S A *101*, 4164-4169.

Capaldi, A.P., Kaplan, T., Liu, Y., Habib, N., Regev, A., Friedman, N., and O'Shea, E.K. (2008). Structure and function of a transcriptional network activated by the MAPK Hog1. Nat Genet *40*, 1300-1306.

Daily, J.P., Scanfeld, D., Pochet, N., Le Roch, K., Plouffe, D., Kamal, M., Sarr, O., Mboup, S., Ndir, O., Wypij, D.*, et al.* (2007). Distinct physiological states of Plasmodium falciparum in malaria-infected patients. Nature *450*, 1091-1095.

Decker, T., Muller, M., and Stockinger, S. (2005). The yin and yang of type I interferon activity in bacterial infection. Nat Rev Immunol *5*, 675-687.

Dillon, S., Agrawal, A., Van Dyke, T., Landreth, G., McCauley, L., Koh, A., Maliszewski, C., Akira, S., and Pulendran, B. (2004). A Toll-like receptor 2 ligand stimulates Th2 responses in vivo, via induction of extracellular signal-regulated kinase mitogen-activated protein kinase and c-Fos in dendritic cells. J Immunol *172*, 4733-4743.

Doyle, S., Vaidya, S., O'Connell, R., Dadgostar, H., Dempsey, P., Wu, T., Rao, G., Sun, R., Haberland, M., Modlin, R.*, et al.* (2002). IRF3 mediates a TLR3/TLR4-specific antiviral gene program. Immunity *17*, 251-263.

Erwin, D.H., and Davidson, E.H. (2009). The evolution of hierarchical gene regulatory networks. Nat Rev Genet *10*, 141-148.

Geiss, G.K., Bumgarner, R.E., Birditt, B., Dahl, T., Dowidar, N., Dunaway, D.L., Fell, H.P., Ferree, S., George, R.D., Grogan, T.*, et al.* (2008). Direct multiplexed measurement of gene expression with color-coded probe pairs. Nat Biotechnol *26*, 317-325.

Hindorff, L.A., Sethupathy, P., Junkins, H.A., Ramos, E.M., Mehta, J.P., Collins, F.S., and Manolio, T.A. (2009). Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc Natl Acad Sci U S A *106*, 9362-9367.

Hu, Z., Killion, P.J., and Iyer, V.R. (2007). Genetic reconstruction of a functional transcriptional regulatory network. Nat Genet *39*, 683-687.

Huang, Q., Liu, D., Majewski, P., Schulte, L.C., Korn, J.M., Young, R.A., Lander, E.S., and Hacohen, N. (2001). The plasticity of dendritic cell responses to pathogens and their components. Science *294*, 870-875.

Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U., and Speed, T.P. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostat *4*, 249-264.

Kagan, J.C., Su, T., Horng, T., Chow, A., Akira, S., and Medzhitov, R. (2008). TRAM couples endocytosis of Toll-like receptor 4 to the induction of interferon-beta. Nat Immunol *9*, 361-368.

Kawai, T., and Akira, S. (2009). The roles of TLRs, RLRs and NLRs in pathogen recognition. Int Immunol *21*, 317-337.

Kawai, T., Sato, S., Ishii, K.J., Coban, C., Hemmi, H., Yamamoto, M., Terai, K., Matsuda, M., Inoue, J., Uematsu, S.*, et al.* (2004). Interferon-alpha induction through Toll-like receptors involves a direct interaction of IRF7 with MyD88 and TRAF6. Nat Immunol *5*, 1061-1068.

Kim, H.D., Shay, T., O'Shea, E.K., and Regev, A. (2009). Transcriptional Regulatory Circuits:Predicting Numbers from Alphabets. Science *325*, 429-32.

Lee, S.I., Dudley, A.M., Drubin, D., Silver, P.A., Krogan, N.J., Pe'er, D., and Koller, D. (2009). Learning a prior on regulatory potential from eQTL data. PLoS Genet *5*, e1000358.

Li, B., Zhou, J., Liu, P., Hu, J., Jin, H., Shimono, Y., Takahashi, M., and Xu, G. (2007). Polycomb protein Cbx4 promotes SUMO modification of de novo DNA methyltransferase Dnmt3a. Biochem J *405*, 369-378.

Litvak, V., Ramsey, S.A., Rust, A.G., Zak, D.E., Kennedy, K.A., Lampano, A.E., Nykter, M., Shmulevich, I., and Aderem, A. (2009). Function of C/EBPdelta in a regulatory circuit that discriminates between transient and persistent TLR4-induced signals. Nat Immunol *10*, 437-443.

Luster, A.D. (2002). The role of chemokines in linking innate and adaptive immunity. Curr Opin Immunol *14*, 129-135.

Mangan, S., and Alon, U. (2003). Structure and function of the feed-forward loop network motif. Proc Natl Acad Sci U S A *100*, 11980-11985.

Matys, V., Kel-Margoulis, O.V., Fricke, E., Liebich, I., Land, S., Barre-Dirrie, A., Reuter, I., Chekmenev, D., Krull, M., Hornischer, K.*, et al.* (2006). TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. Nucleic Acids Res *34*, D108-110.

Moffat, J., Grueneberg, D.A., Yang, X., Kim, S.Y., Kloepfer, A.M., Hinkle, G., Piqani, B., Eisenhaure, T.M., Luo, B., Grenier, J.K.*, et al.* (2006). A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. Cell *124*, 1283-1298.

Pe'er, D., Regev, A., and Tanay, A. (2002). Minreg: inferring an active regulator set. Bioinformatics *18 Suppl 1*, S258-267.

Ramsey, S.A., Klemm, S.L., Zak, D.E., Kennedy, K.A., Thorsson, V., Li, B., Gilchrist, M., Gold, E.S., Johnson, C.D., Litvak, V.*, et al.* (2008). Uncovering a macrophage transcriptional program by integrating evidence from motif scanning and expression dynamics. PLoS Comput Biol *4*, e1000021.

Raychaudhuri, S., Plenge, R.M., Rossin, E.J., Ng, A.C., Purcell, S.M., Sklar, P., Scolnick, E.M., Xavier, R.J., Altshuler, D., and Daly, M.J. (2009). Identifying relationships among genomic disease regions: predicting genes at pathogenic SNP associations and rare deletions. PLoS Genet *5*, e1000534.

Sandelin, A., Alkema, W., Engstrom, P., Wasserman, W.W., and Lenhard, B. (2004). JASPAR: an open-access database for eukaryotic transcription factor binding profiles. Nucleic Acids Res *32*, D91-94.

Segal, E., Shapira, M., Regev, A., Pe'er, D., Botstein, D., Koller, D., and Friedman, N. (2003). Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data. Nat Genet *34*, 166-176.

Suzuki, H., Forrest, A.R., van Nimwegen, E., Daub, C.O., Balwierz, P.J., Irvine, K.M., Lassmann, T., Ravasi, T., Hasegawa, Y., de Hoon, M.J.*, et al.* (2009). The transcriptional network that controls growth arrest and differentiation in a human myeloid leukemia cell line. Nat Genet *41*, 553-562.

Wang, X., Arai, S., Song, X., Reichart, D., Du, K., Pascual, G., Tempst, P., Rosenfeld, M.G., Glass, C.K., and Kurokawa, R. (2008). Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. Nature *454*, 126-130.

Workman, C.T., Mak, H.C., McCuine, S., Tagne, J.B., Agarwal, M., Ozier, O., Begley, T.J., Samson, L.D., and Ideker, T. (2006). A systems approach to mapping DNA damage response pathways. Science *312*, 1054-1059.

Zou, H., and Hastie, T. (2005). Regularization and variable selection via the Elastic Net. Journal of the Royal Statistical Society B *67*, 301-320.

Chapter 3 – Systematic Discovery of TLR Signaling Components Delineates Viral-Sensing Circuits

3.1. Author Contributions

Reference (reproduced with permission from *Cell*):

Chevrier, N., Mertins, P., Artyomov, M.N., Shalek, A.K., Iannacone, M., Ciaccio, M.F., Gat-Viks, I., Tonti, E., DeGrace, M.M., Clauser, K.R., Garber, M., Eisenhaure, T.M., Yosef, N., Robinson, J., Sutton, A., Andersen, M.S., Root, D.E., von Andrian, U., Jones, R.B., Park, H., Carr, S.A., Regev, A.*, Amit, I.*, Hacohen, N.* (2011). Systematic discovery of TLR signaling components delineates viral-sensing circuits. Cell *147*, 853-867. (* equal contributions)

- Nicolas Chevrier conceived the study, designed all experiments and performed almost all of them (except confocal microscopy, *in vivo* infections, and primary mouse lung fibroblast infections), analyzed and interpreted data, and wrote the manuscript.
- Philipp Mertins optimized the phosphoproteomics protocol, assisted with execution of proteomics experiments, operated the mass spectrometer, and processed raw data files.
- Maxim N. Artyomov contributed to the statistical analysis of the NanoString data.
- Alex K. Shalek provided nanowires for optimization experiments, and performed the confocal microscopy experiments.
- Matteo Iannacone performed *in vivo* VSV infection and tissue harvest.
- Mark F. Ciaccio operated the microwestern array platform and assisted with data analysis.
- Irit Gat-Viks assisted with microarray data analysis.

- Elena Tonti assisted with *in vivo* VSV infection and tissue harvest.
- Marciela M. DeGrace performed influenza infection experiments in primary mouse lung fibroblasts.
- Karl R. Clauser assisted with phosphotyrosine-based phosphoproteomics experiments, and with computational analysis of the mass spectrometry raw data sets.
- Manuel Garber assisted with NanoString data analysis.
- Thomas M. Eisenhaure assisted with RNAi experiments.
- Nir Yosef assisted with experimental design of the drug/phosphoproteomics experiments.
- Jacob Robinson, Amy Sutton, and Mette S. Andersen assisted with nanowire experiments.
- David E. Root supervised RNAi experiments and provided reagents.
- Ulrich von Andrian supervised the *in vivo* infection experiments.
- Richard B. Jones supervised the microwestern array-based experiments.
- Hongkun Park supervised the nanowire-based experiments.
- Steven A. Carr supervised the mass spectrometry-based work.
- Aviv Regev supervised the computational aspect of the project and manuscript writing.
- Ido Amit supervised the experimental aspect of the project and manuscript writing.
- Nir Hacohen supervised the entire project and manuscript writing.

3.2. Introduction

Signaling networks function to detect environmental changes and to trigger appropriate responses. Defects in the wiring of these networks can contribute to diseases. For example, Toll-like receptors (TLRs) in innate immune dendritic cells (DCs) sense microbial components and trigger signaling pathways critical for host defense (Takeuchi and Akira, 2010). Genetic

defects in components of the TLR network (*e.g.*, *Tnfaip3*, *Irf5*, *Tlr4*) have been linked to human inflammatory diseases (Hennessy et al., 2010).

Deciphering signaling networks is a challenging task. First, despite extensive studies, many components of key networks remain unknown, limiting our ability to understand the role of genes genetically associated with disease (Xavier and Rioux, 2008). Second, there are few systematic approaches to determine the function of a new component and place it accurately within pathways. Third, it is hard to compare and connect signaling nodes that have been studied in disparate systems and with different readouts. Filling in such gaps is critical for understanding signal processing in cells and for manipulating clinically relevant pathways.

In a recent study (Amit et al., 2009), we developed a perturbational strategy to decipher transcriptional circuits and applied it to the TLR system of immune DCs (see Chapter 2). In this approach, we identified transcriptional regulators based on changes in their mRNA levels following TLR stimulation, perturbed each regulator with an shRNA, and measured the effect using a 118-gene signature. This allowed us to identify functional transcriptional regulators, and to associate them with specific targets and with the global responses they control. For example, we discovered that a host of cell proliferation-associated factors, such as *Rbl1*, *Rb*, *Myc*, *Jun*, and *E2fs* have been co-opted to control the antiviral transcriptional program in non-dividing terminal DCs.

Here, we adapted and expanded this approach to develop a strategy for the discovery and validation of signaling components (**Figure 3.1**). We relied on the fact that in many responses, including pathogen sensing, there often are transcriptional feedbacks, whereby a signaling circuit also regulates the transcript levels of genes encoding some, but not all, of its components (Amit et al., 2007; Fraser and Germain, 2009; Segal et al., 2003). We therefore initially followed

the same strategy as that used for transcriptional regulators to identify candidate signaling molecules (from mRNA profiles), perturb them (using shRNAs), and measure the effect on a representative gene signature. Furthermore, to expand the pathway's scope to additional components whose mRNA levels may be unchanged, we also used phosphoproteomics following perturbation of the initially discovered signaling molecules.

We applied this iterative approach to study the TLR response, discovering 19 new functional components of TLR signaling pathways, including a new major arm in antiviral signaling. This arm is centered on *Plk2* and *4*, two Polo-like kinases that play a prominent and previously unknown role in all well-described host antiviral pathways, both *in vitro* and *in vivo*. Using phosphoproteomics, followed by RNAi, we identified eleven additional members of a Plk-dependent, antiviral signaling module. These include *Tnfaip2*, which was recently genetically associated with two autoimmune disorders but whose role was unknown. Our work thus establishes an effective strategy to systematically assign functions to newly identified signaling components of a network, and to point towards potential therapeutic targets.

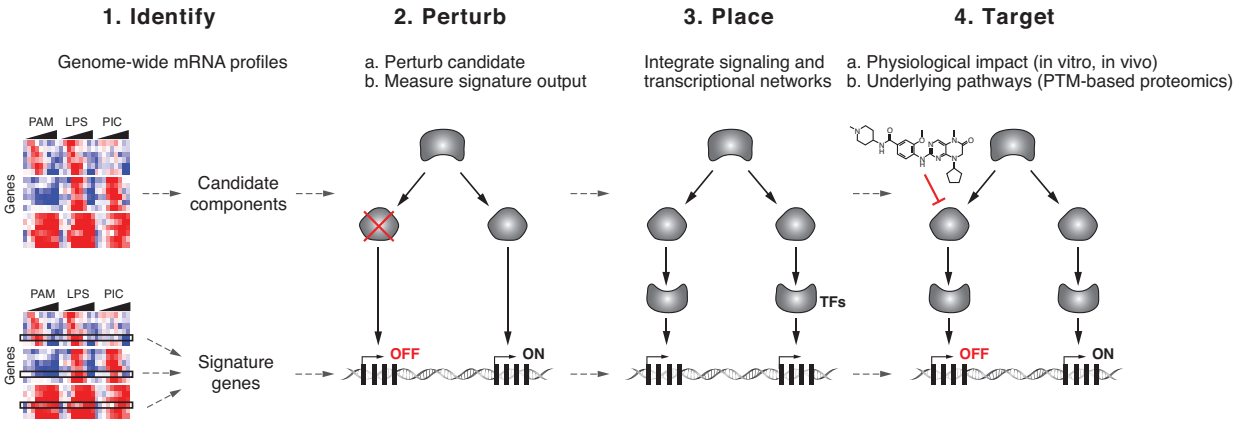

Figure 3.1. A Systematic Approach to Dissect Signaling Pathways

Shown is a schematic depicting the strategy consisting of 4 steps (from left to right): (1) extract both candidate signaling regulators and signature genes; (2) perturb each candidate with shRNAs and measure the effect on the expression of signature genes; (3) compare perturbation profiles of signaling and transcriptional regulators to start assembling pathways; (4) use small molecule targeting of signaling nodes of interest to a) evaluate the physiological relevance of new signaling node, and b) identify underlying pathways by discovering downstream effector molecules.

3.3. Transcripts for Signaling Components Are Regulated upon TLR Stimulation

To analyze pathogen-sensing pathways, we studied DCs stimulated with agonists for TLR2, 3, or 4. These TLRs activate transcriptional programs referred to, here, as "inflammatory" (TLR2), "antiviral" (TLR3), or both (TLR4) (**Figure 3.2A**) (Amit et al., 2009; Doyle et al., 2002). The signaling pathways mediating these responses rely on both shared and unique components. For example, both antiviral and inflammatory pathways activate the MAPK and NF-κB signaling cascades, whereas only antiviral pathways activate the TBK1/IKK-ε/IRFs axis that leads to type I interferon (IFNs) production (**Figure 3.2A**) (Takeuchi and Akira, 2010).

We chose candidate components in the TLR pathways by identifying genes encoding putative signaling proteins whose expression levels change following TLR stimulation (**Figure 3.2B**). We used genome-wide mRNA expression profiles, previously measured (Amit et al., 2009) at 10 time points along 24 hours following stimulation of bone marrow-derived DCs (BMDCs) with lipopolysaccharide (LPS; TLR4 agonist), polyinosinic:polycytidylic acid (poly(I:C); recognized by TLR3 and the cytosolic viral sensor MDA-5), or Pam3CSK4 (PAM; TLR2 agonist) (**Figure 3.2B**). 280 genes annotated as known or putative signaling components in the genome were differentially expressed following stimulation, including 115 kinases, 69 phosphatases, and 96 adaptors, scaffolds and other signaling regulators (**Figure 3.2B, and Experimental Procedures**). These 280 genes were enriched for canonical pathways of the TLR network such as MAP kinase ($P < 1.22\times10^{-15}$, overlap 25/87, hypergeometric test), TLR (*e.g.*, *Myd88*, *Traf6*, *Irak4*, *Tbk1*; $P < 8.43\times10^{-12}$, 21/86), and PI3K ($P < 2.58\times10^{-8}$, 11/33) pathways, as well as the PYK2 pathway ($P < 3.12\times10^{-10}$, overlap 12/29), which was recently associated with the TLR system (Wang et al., 2010). Overall, 94 of the 280 genes (33%) were previously associated with the TLR network in the literature, supporting the validity of our strategy. The remaining 186 (67%) genes represent novel candidate components of the TLR system. Of these 186, we

selected 23 novel candidates for further experiments based on their strong differential expression and to proportionally represent the five main induced expression clusters (**Figure 3.2B and 3.2C**). Six canonical TLR components (*Myd88*, *Mapk9*, *Tbk1*, *Ikbke*, *Tank*, and *Map3k7*) were selected as positive controls to benchmark our approach (**Figure 3.2A and 3.2D**).

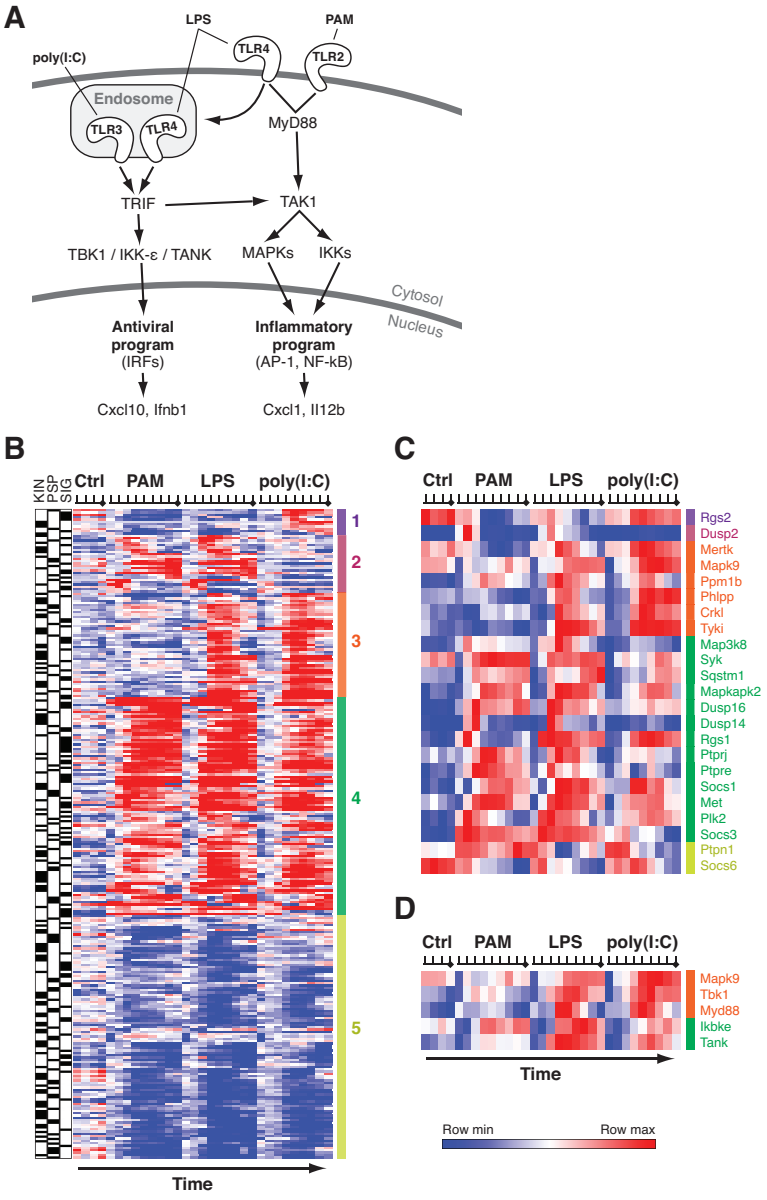

Figure 3.2. mRNAs of Signaling Components Are Differentially Regulated upon TLR Stimulation
(A) Schematic depicting simplified pathways triggered by TLR2, 3, and 4 (Reviewed in Takeuchi and Akira, 2010).

Figure 3.2 (Continued)
(B) mRNA expression profiles of differentially expressed signaling genes. Shown are expression profiles for 280 differentially expressed signaling genes (rows) across different time points (columns): a control time course (no stimulation, Ctrl) and following stimulations with Pam3CSK4 (TLR2 agonist, PAM), lipopolysaccharide (TLR4 agonist, LPS), and poly(I:C) (TLR3 agonist). Tick marks: time point post-stimulation (0.5, 1, 2, 4, 6, 8, 12, 16, 24 hours). Shown genes had at least a 1.7 fold increase or decrease in expression compared to pre-stimulation levels in both duplicates of at least one time point. The three leftmost columns indicate kinase (KIN), phosphatase (PSP), and signaling (SIG) regulator functions (black bars). Values from duplicate arrays were collapsed and gene expression profiles were hierarchically clustered: the rightmost color-coded column indicates the 5 major expression clusters.
(C and D) mRNA expression profiles of candidate (C) and canonical (D) TLR signaling regulators that were selected for subsequent experiments. The color-coding of the gene names highlight the corresponding expression cluster from the complete matrix of 280 signaling genes from (B).

3.4. A Perturbation Strategy Places Uncharacterized Signaling Components within the Antiviral and Inflammatory Pathways

We perturbed each of the 6 positive controls and 17 of the 23 novel candidates in DCs with lentivirus-delivered shRNAs (**Experimental Procedures**), stimulated the cells with LPS, and measured the effect on the mRNA levels of 118 signature genes, plus 10 control genes, at 6 hours post-stimulation (**Figure 3.3A**). (Six of the 23 initial candidates were not studied further due to poor knockdown efficiency.)

We have previously shown that there is no significant alteration in function in DCs infected with shRNA-encoding lentiviruses and selected with puromycin for 4 days during GM-CSF-induced differentiation (Amit et al., 2009). The stimulus, time point, and gene signature were defined in our previous study as representatives of both the inflammatory and the antiviral programs (**Experimental Procedures**) (Amit et al., 2009).

We used a multiplex mRNA counting methodology to quantify the transcript levels of these signature genes, and determined statistically significant changes in signature transcripts by relying on both control genes and control shRNAs (**Experimental Procedures**) (Geiss et al., 2008). Finally, we associated signaling molecules and downstream transcriptional regulators

that may act in the same pathway by comparing the perturbational profiles of the 23 signaling molecules (6 canonical and 17 candidates) to each other and to those of the 123 transcription regulators previously tested (**Figure 3.3 and Figure 3.4**) (Amit et al., 2009).

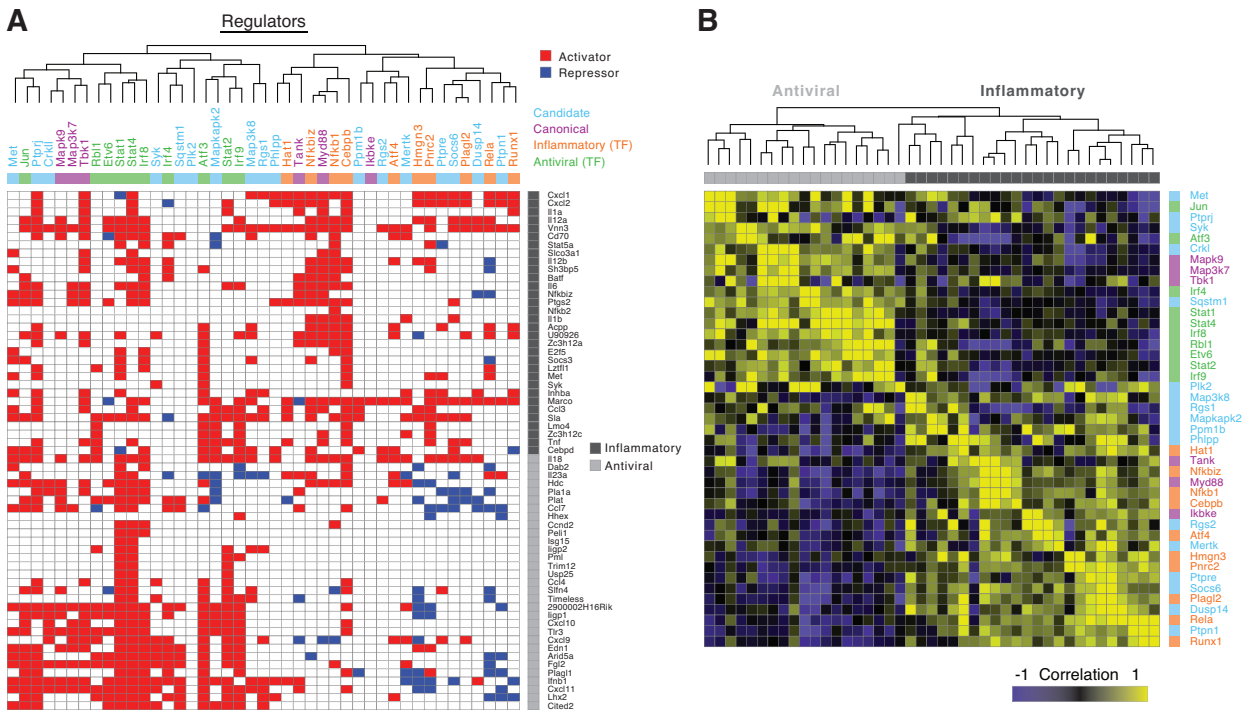

Figure 3.3. A Perturbation Strategy Assigns Function to Signaling Components within the TLR Pathways

(A) Perturbation profiles of six canonical (purple) and 17 candidate (light blue) signaling components, and 20 core TLR transcriptional regulators belonging to the inflammatory (orange) and the antiviral (green) programs. Shown are the perturbed regulators (columns) and their statistically significant effects (False discovery rate, FDR < 0.02) on each of the 118 TLR signature genes (rows). Red: significant activating relation (target gene expression decreased following perturbation); blue: significant repressing relation (target gene expression increased following perturbation); white: no significant effect. The column on the right indicates whether signature genes belong to the antiviral (light grey) or the inflammatory (dark grey) programs.

(B) Functional characterization based on similarity of perturbation profiles. Shown is a correlation (Pearson) matrix of the perturbation profiles from A. Yellow: positive correlation; purple: negative correlation; black: no correlation.

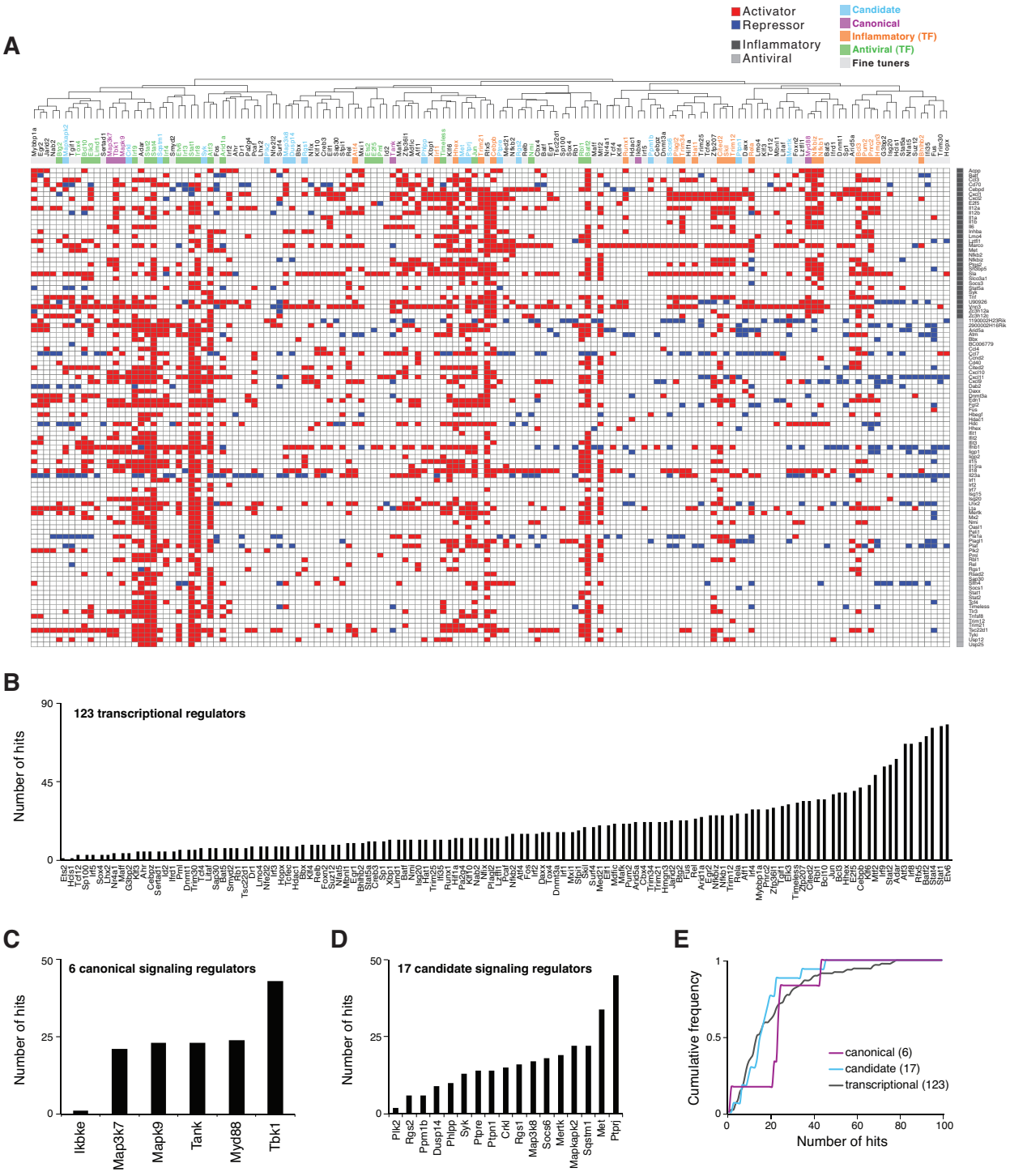

Figure 3.4. Perturbations of Signaling and Transcriptional Regulators Have Similar Effects on the TLR Signature Genes

(A) Perturbation profiles of 6 canonical (purple) and 17 candidate (light blue) signaling regulators, and 123 transcriptional regulators (TF) partitioned into regulators of the inflammatory (orange) and antiviral (green) programs, and fine tuners (grey), as previously defined in Amit et al., 2009. Shown are the perturbed regulators (columns) and their statistically significant effects (False discovery rate, FDR < 2%) on each of the 118 TLR signature genes (rows). Red: significant activating relation (target gene expression decreased following perturbation); blue: significant repressing relation (target gene expression increased

Figure 3.4 (Continued)
following perturbation); white: no significant effect. The column on the right indicates whether signature genes belong to the antiviral (light grey) or the inflammatory (dark grey) programs.
(B-D) Shown are the numbers of signature gene hits (Y axis) significantly affected by knockdown of each regulator (X axis) for the regulator categories shown in A: 123 transcriptional (B) and 6 previously known (C) and 17 candidate (D) signaling regulators.
(E) Candidate signaling regulators affect a similar number of 'signature' genes compared to transcriptional regulators. Shown is the cumulative distribution of the number of hits for the regulators shown in B-D.

Perturbing five of the six positive control signaling molecules strongly affected the expression of TLR signature genes in a manner consistent with their known roles (**Figure 3.3A**), thus validating our approach. For example, perturbing the known inflammatory adaptor *Myd88* specifically abrogated the transcription of inflammatory genes (*e.g.*, *Cxcl1*, *Il1a*, *Il1b*, *Ptgs2*, *Tnf*; **Figure 3.3A**), an effect similar to that of perturbations of downstream inflammatory transcription factors (*e.g.*, *Nfkb1*, *Nfkbiz*; **Figure 3.3B**). In addition, *Tank* acted as a negative regulator of a subset of antiviral genes (**Figure 3.3A**), consistent with previous observations (Kawagoe et al., 2009), and *Tbk1* knockdown affected both antiviral and inflammatory outputs (**Figure 3.3A**), consistent with recent findings that *Tbk1* regulates NF-κB complexes (Barbie et al., 2009; Chien et al., 2006). Notably, *Ikbke* (IKK-ε) knockdown did not affect our gene signature, consistently with previous observations that IKK-$\varepsilon^{-/-}$ DCs respond normally to LPS and viral challenges (Matsui et al., 2006). Thus, IKK-ε may either not be functional or be redundant in our system.

All of the 17 candidate signaling molecules tested, except *Plk2* (discussed below), affected at least 6 targets in the 118-gene signature (on average, 16.6 targets ± 10.4 SD), and 12 affected more than 10% of the targets (**Figure 3.4A and 3.4D**). Notably, perturbations of these 17 candidates did not affect BMDC differentiation (88.3% ± 6.8 SD of $CD11c^+$ cells). These effects on TLR signature gene expression are comparable to those observed for known signaling molecules and transcriptional regulators in this system (**Figure 3.4B-E**). For example, the receptor tyrosine kinase *Met*, not previously associated with TLR signaling, affected a number of signature genes similar to *Tbk1* (**Figure 3.4C and 3.4D**), in both the inflammatory and antiviral

programs (**Figure 3.3A**). Conversely, both the phosphatase *Ptpre* and the adaptor *Socs6* acted as positive regulators of the inflammatory program, while negatively regulating a subset of antiviral genes (**Figure 3.3A**). Of the 17 candidates tested when we originally conducted this targeted screen, 10 have subsequently been reported by other groups to be functional in the TLR system, providing an independent confirmation of our results. For example, in our experiments, *Map3k8* knockdown affected both inflammatory and antiviral target genes (**Figure 3.3A**), consistent with its reported role in the TLR pathways based on phenotypic analysis of *Sluggish* mice (Xiao et al., 2009).

Our approach identified both primary (*e.g.*, *Myd88*) and secondary (*e.g.*, *Stat1*) mediators of TLR responses. While secondary mediators are not part of the initial intracellular signaling cascade, they are important physiological components of the TLR response and their perturbation leads to the same phenotypic outcome as primary components. For example, we re-discovered a member of the TAM family of receptor tyrosine kinases, *Mertk*, which acts secondarily as a positive regulator of part of the inflammatory program, and as a negative regulator of several antiviral genes, including *Ifnb1* (**Figure 3.3A**). This is consistent with the reported inhibitory roles of TAM family members in the IFN pathway (Rothlin et al., 2007).

3.5. Crkl Modulates JNK-mediated Antiviral Signaling in the TLR Network

Among the 17 candidate signaling proteins, perturbation of the tyrosine kinase adaptor *Crkl* decreased expression of a relatively large fraction (13%) of the signature genes (**Figure 3.3A and Figure 3.4D**), especially antiviral ones. *Crkl* is a known component of several signaling pathways, including early lymphocyte activation (Birge et al., 2009), but has not been previously associated with the TLR network. *Crkl*'s perturbation profile closely resembled those of known antiviral regulators, most notably JNK2 (*Mapk9*; (Chu et al., 1999) (**Figure 3.3A and 3.5A**).

Indeed, when *Crkl*$^{-/-}$ DCs were stimulated with LPS, the expression of antiviral cytokines (*Cxcl10*, *Ifnb1*) was strongly and specifically reduced (**Figure 3.5B**, left and middle), whereas that of an inflammatory cytokine (*Cxcl1*) remained unaffected (**Figure 3.5B**, right).

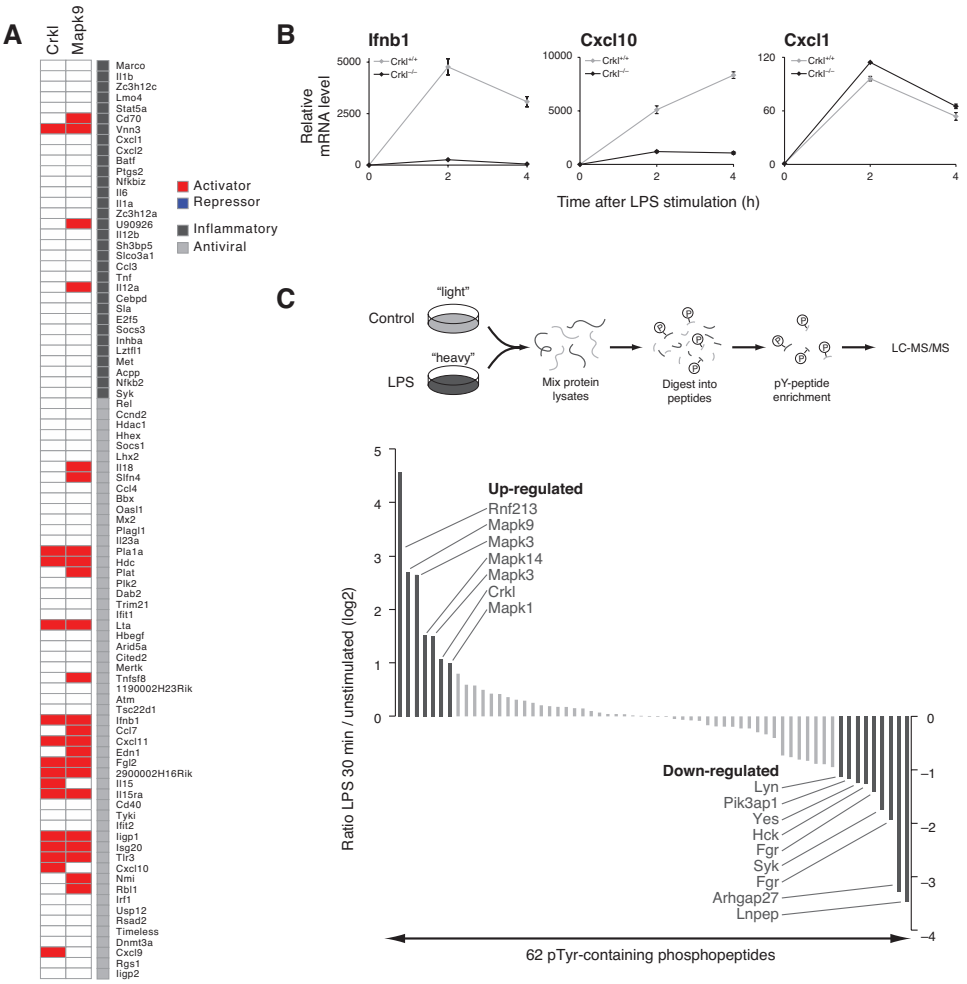

Figure 3.5. *Crkl* Adaptor Functions in the Antiviral Arm of TLR4 Signaling

(A) Comparison of *Crkl* and *Mapk9* knockdown profiles. Shown are the effects of *Crkl* and *Mapk9* perturbation (columns) on the 118 signature genes (rows). Data was extracted and is presented as in Figure 3.3A.

(B) Inhibition of transcription of mRNAs of antiviral cytokines in *Crkl*$^{-/-}$ BMDCs. Shown are mRNA levels (qPCR; relative to t = 0) for *Ifnb1* (left), *Cxcl10* (middle) and *Cxcl1* (right) in three replicates per time point. Error bars represent the standard error of the mean (n = 3 mice).

(C) Crkl phosphorylation is induced following LPS stimulation. Top: Schematic depiction of experimental workflow. From left: Protein lysates from unstimulated (Control) and LPS-treated BMDCs grown in "light" and "heavy" SILAC medium were mixed (1:1) and digested into peptides with trypsin before phospho-tyrosine (pY) peptide enrichment by immunoprecipitation, and LC-MS/MS analysis. Bottom: Shown are the differential phosphorylation levels (log 2 ratios, Y axis) of all 62 phosphopeptides identified and quantified by LC-MS/MS (X axis). Black: peptides with more than 2 fold differential expression (left: induced; right: repressed).

To test whether *Crkl* is a primary component of the TLR pathway, we determined whether Crkl phosphorylation is rapidly modified after TLR signaling initiation. We used SILAC-based quantitative phosphoproteomics that led to the identification and quantification of 62 phospho-tyrosine (pTyr)-containing peptides from BMDCs stimulated with LPS for 30 minutes (**Figure 3.5C and Experimental Procedures**) (Ong et al., 2002). Of these 62 phosphopeptides that were identified and quantified, 7 and 9 were significantly up- or down-regulated, respectively (**Figure 3.5C**). A pTyr-containing peptide derived from Crkl (Y132) – one of the top-six induced phosphopeptides – was induced 2.1 fold (**Figure 3.5C**). This observation corroborates the hypothesis that Crkl is active immediately downstream of TLR4 signaling.

Several lines of evidence suggest that Crkl acts through Jnk2 (*Mapk9*) signaling. First, following LPS stimulation, the MAP kinase Jnk2 (*Mapk9*) is co-regulated at the phosphorylation level with Crkl (**Figure 3.5C**). Second, the Crk adaptor family – including CrkI, CrkII, and Crkl – has been previously reported to modulate Jnk activity in growth factor and IFN signaling (Birge et al., 2009; Hrincius et al., 2010). Third, the perturbation profiles of *Mapk9* and *Crkl* are strikingly similar (**Figure 3.5A**). Taken together, these observations suggest that Crkl modulates Jnk-mediated antiviral signaling in the TLR4 pathway.

3.6. Polo-like Kinases Are Critical Activators of the Antiviral Program

To discover potential drug targets among our 17 candidates, we next focused on Polo-like kinase (Plk) 2, a well-known cell cycle regulator and drug target (Strebhardt, 2010), which was strongly induced by a viral-like stimulus. The roles of Plks in non-dividing, differentiated cells are poorly defined (Archambault and Glover, 2009; Strebhardt, 2010). We have previously shown that transcriptional regulators of cell cycle processes (*e.g.*, *Rbl1*, *Rb*, *Myc*, *Jun*, *E2fs*) are co-

opted to function in the antiviral responses in DCs (Amit et al., 2009). We therefore hypothesized that *Plk2* might be an upstream signaling regulator of the antiviral program.

Contrary to our hypothesis, neither knockdown (**Figure 3.3A**) nor knockout (Inglis et al., 2009) of *Plk2* in BMDCs had any effect on the TLR response. We speculated that this could be the result of redundancy with another Plk, since Plk4 mRNA is induced in DCs with a similar pattern to Plk2 (**Figure 3.6A**), albeit at a lower amplitude (and thus was not included in the 17 initial candidates). Interestingly, functional redundancy between *Plk2* and *4* has been suggested to account for the viability of *Plk2*-deficient mice (Strebhardt, 2010), and *Plk2* and *4* have been reported to function together in centriole duplication (Chang et al., 2010; Cizmecioglu et al., 2008).

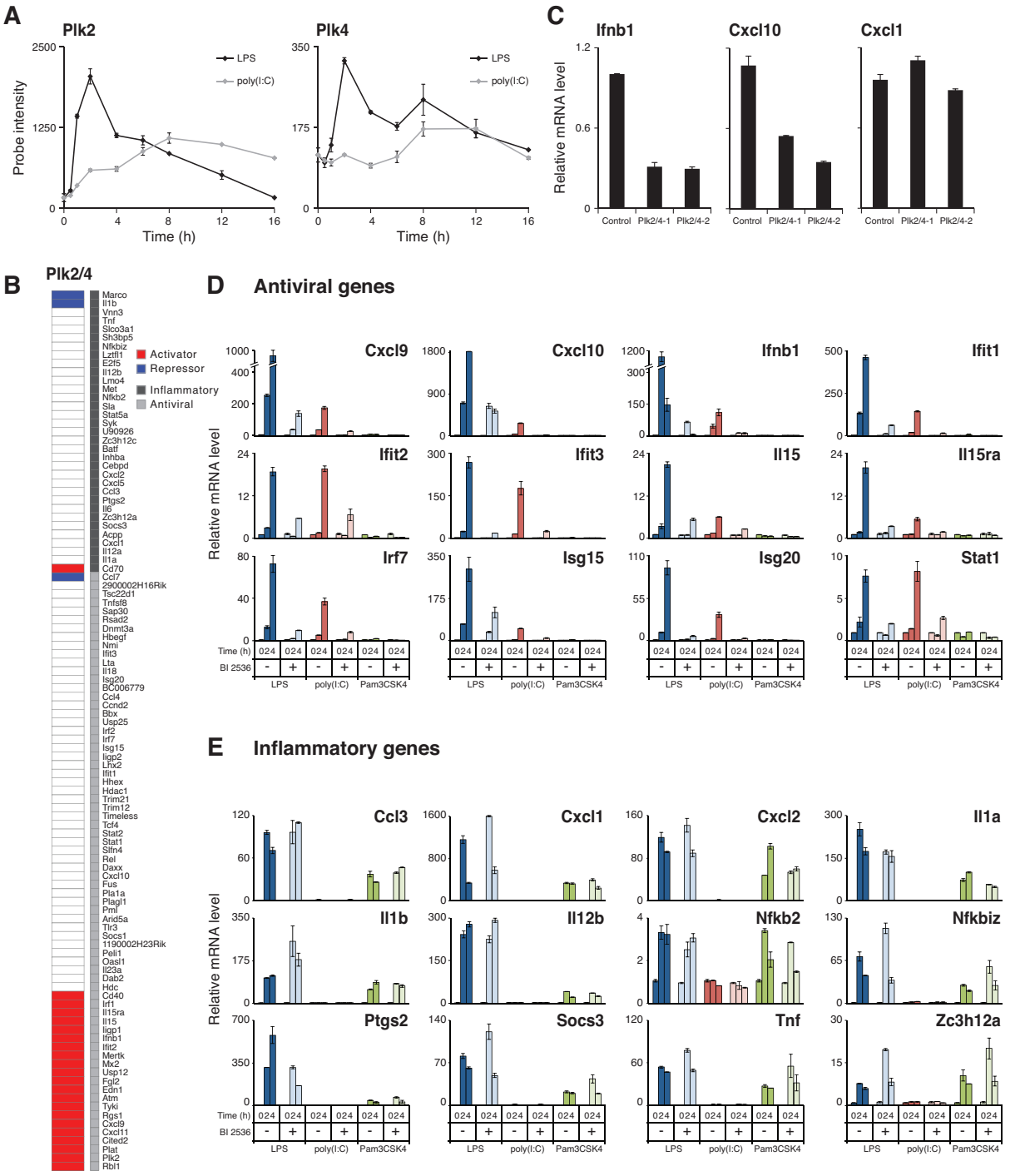

Figure 3.6. *Plk2* and *4* Regulate the Antiviral Program

(A) Similarity of *Plk2* and *4* mRNA expression profiles. Shown are mRNA levels (from **Figure 3.2B**) of *Plk2* (left) and *4* (right) following stimulation with LPS (black) or poly(I:C) (grey).

(B) Double knockdown of *Plk2* and *4* represses the antiviral signature. Shown are significant changes in expression of TLR signature genes (rows) following double knockdown of *Plk2* and *4*. Red and blue mark significant hits as in **Figure 3.3**, only for genes where the effect was consistent between the two independent combinations of shRNAs.

(C) Double knockdown of *Plk2* and *4* represses antiviral cytokine mRNAs. Shown are expression levels (qPCR) relative to control shRNAs (Control) for two antiviral cytokines (*Ifnb1* and *Cxcl10*) and for an

Figure 3.6 (Continued)
inflammatory cytokine (*Cxcl1*), following LPS stimulation in BMDCs using two independent combinations of shRNAs (Plk2/4-1, Plk2/4-2). Three replicates for each experiment; error bars are the standard error of the mean.
(D and E) BI 2536 specifically abrogates transcription of antiviral genes without affecting inflammatory genes following stimulation with LPS, poly(I:C), or Pam3CSK4. Shown are mRNA levels (qPCR; relative to t = 0) for 12 indicated antiviral (D) and 12 inflammatory (E) genes in BMDCs treated with BI 2536 (1 μM; dark color bars) or DMSO vehicle (light color bars) and stimulated for 0, 2 or 4 h with LPS (dark and light blue), poly(I:C) (dark and light red), or Pam3CSK4 (dark and light green). Three replicates in each experiment; error bars are the standard error of the mean.

Consistent with this revised hypothesis, double knockdown of *Plk2* and *4* – by infecting cells with two different shRNA-carrying viruses, one shRNA targeting each gene (**Figure 3.7B)** – led to a significant and specific decrease in expression of 21 antiviral response genes (**Figure 3.6B**). For example, the antiviral cytokines *Ifnb1* and *Cxcl10* mRNAs were decreased, whereas the expression of the inflammatory gene *Cxcl1* and almost all inflammatory signature genes remained unaffected (**Figure 3.6C**). This specific response was observed using two independent mixes of shRNAs, in each case targeting Plks with a different pair of shPlk2/shPlk4 (**Figure 3.7B and Experimental Procedures**). Two recent reports suggested a role for *Plk1* alone as a negative regulator of MAVS (Vitour et al., 2009) and NF-κB (Zhang et al., 2010) in cell lines. In BMDCs, however, both *Plk1* and *Plk3* are expressed at very low levels (100-fold lower than *Plk2*), and their knockdown did not affect the TLR transcriptional response (**Figure 3.7C**). Notably, we did not observe a reduction in DC viability when cells were transduced with lentiviral shRNAs targeting *Plk1*, *2*, *3* or *4* individually, or *Plk2* and *4* together (based on mRNA levels of control genes). We thus conclude that in differentiated BMDCs, *Plk2* and *4*, but likely not *Plk1* or *3*, are critical to the regulation of antiviral but not cell cycle pathways.

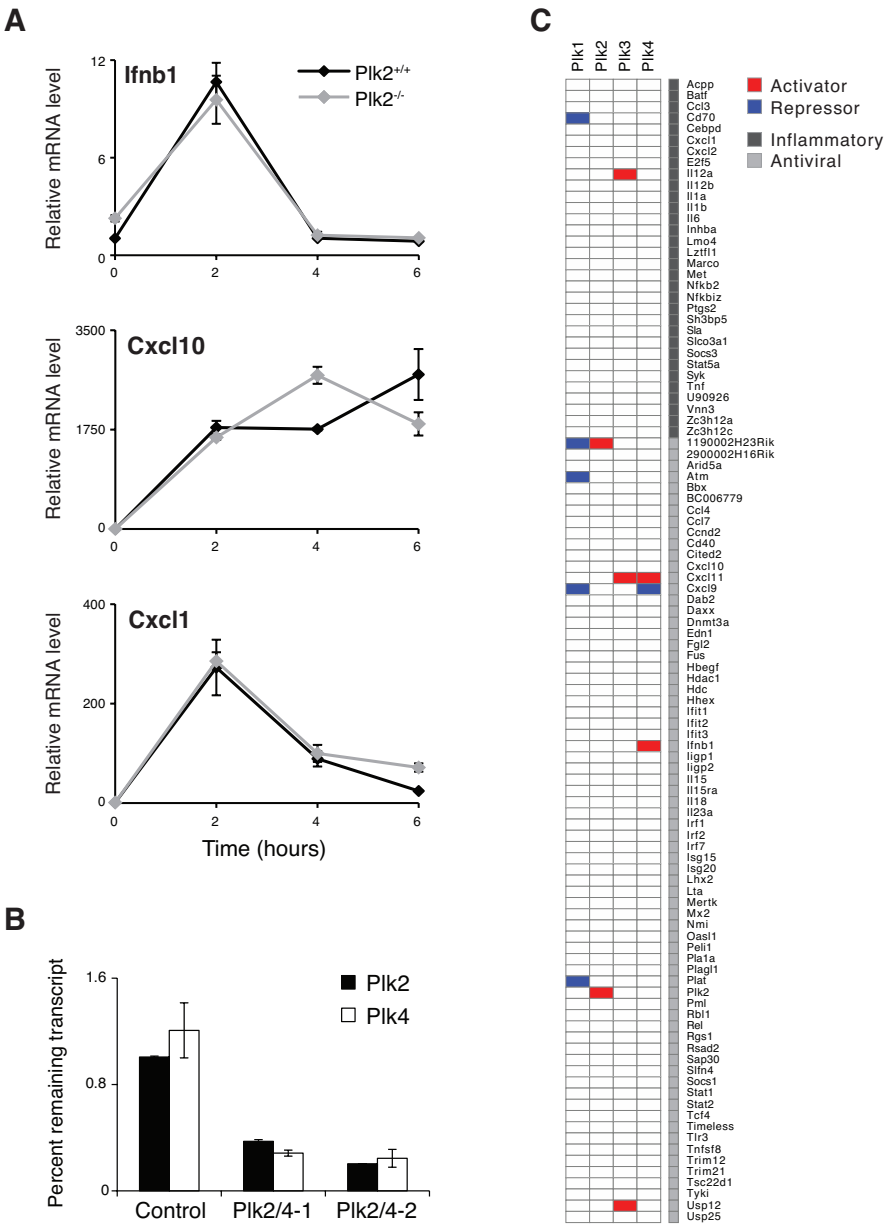

Figure 3.7. Individual Perturbation of Plk Family Members Does Not Affect TLR Output Gene Expression in DCs

(A) *Plk2*-deficient BMDCs respond to LPS similarly to wild-type cells. Shown are mRNA levels (qPCR; relative to t = 0) for *Ifnb1* (left), *Cxcl10* (middle) and *Cxcl1* (right) in three replicates per time point. Error bars represent the standard error of the mean.

(B) Combinatorial knockdown levels of *Plk2* and *4* in BMDCs. Shown are mRNA levels (qPCR), relative to non-targeting shRNAs (Control), of *Plk2* and *4* in BMDCs using two independent combinations of shRNAs (Plk2/4-1 and -2). Three replicates in each experiment; error bars represent the standard error of the mean.

(C) Perturbations of individual Plk family members do not affect TLR signature genes. Shown are the perturbed Plks (columns) and their statistically significant effects (FDR < 2%) on each 118 TLR signature genes (rows). Red: significant activating relation (target gene expression decreased following perturbation); blue: significant repressing relation (target gene expression increased following perturbation); white: no significant effect. The column on the right indicates whether signature genes belong to the antiviral (light grey) or the inflammatory (dark grey) programs.

3.7. A Small Molecule Inhibitor of Plks Represses Antiviral Gene Expression and IRF3 Translocation in DCs

We next manipulated the antiviral responses in DCs using BI 2536, a commercial pan-specific Plk small molecule inhibitor (Steegmaier et al., 2007). First, we compared genome-wide mRNA profiles measured in LPS- or poly(I:C)-stimulated DCs that were pre-treated with either BI 2536 or DMSO vehicle (**Experimental Procedures**). Most genes repressed by BI 2536 treatment but induced by either stimulus in the presence of vehicle, belonged to the antiviral program (99/193 genes in response to poly(I:C), $P < 1\times10^{-71}$, hypergeometric test; 67/194 in response to LPS). The 311 unique LPS- and/or poly(I:C)-induced genes that are repressed by BI 2536, are significantly enriched for genes in canonical pathways related to cytokine signaling (*e.g.*, IL-10, type I IFNs, IL-1), TLR signaling, and DC signaling, and for GO processes related to defense and immune responses (**Figure 3.8A**). This suggests a specific action of BI 2536 on the TLR system of DCs.

To further investigate the Plks' specificity of action, we measured the expression of 12 well-studied antiviral and 12 well-studied inflammatory response genes in DCs stimulated with LPS, poly(I:C), or Pam3CSK4 in the presence and absence of BI 2536 (**Figure 3.6D**). All antiviral genes induced by LPS or poly(I:C) were strongly inhibited upon BI 2536 treatment, whereas inflammatory gene expression remained largely unaffected upon LPS or Pam3CSK4 stimulation, in agreement with the microarray data (**Figure 3.6D**).

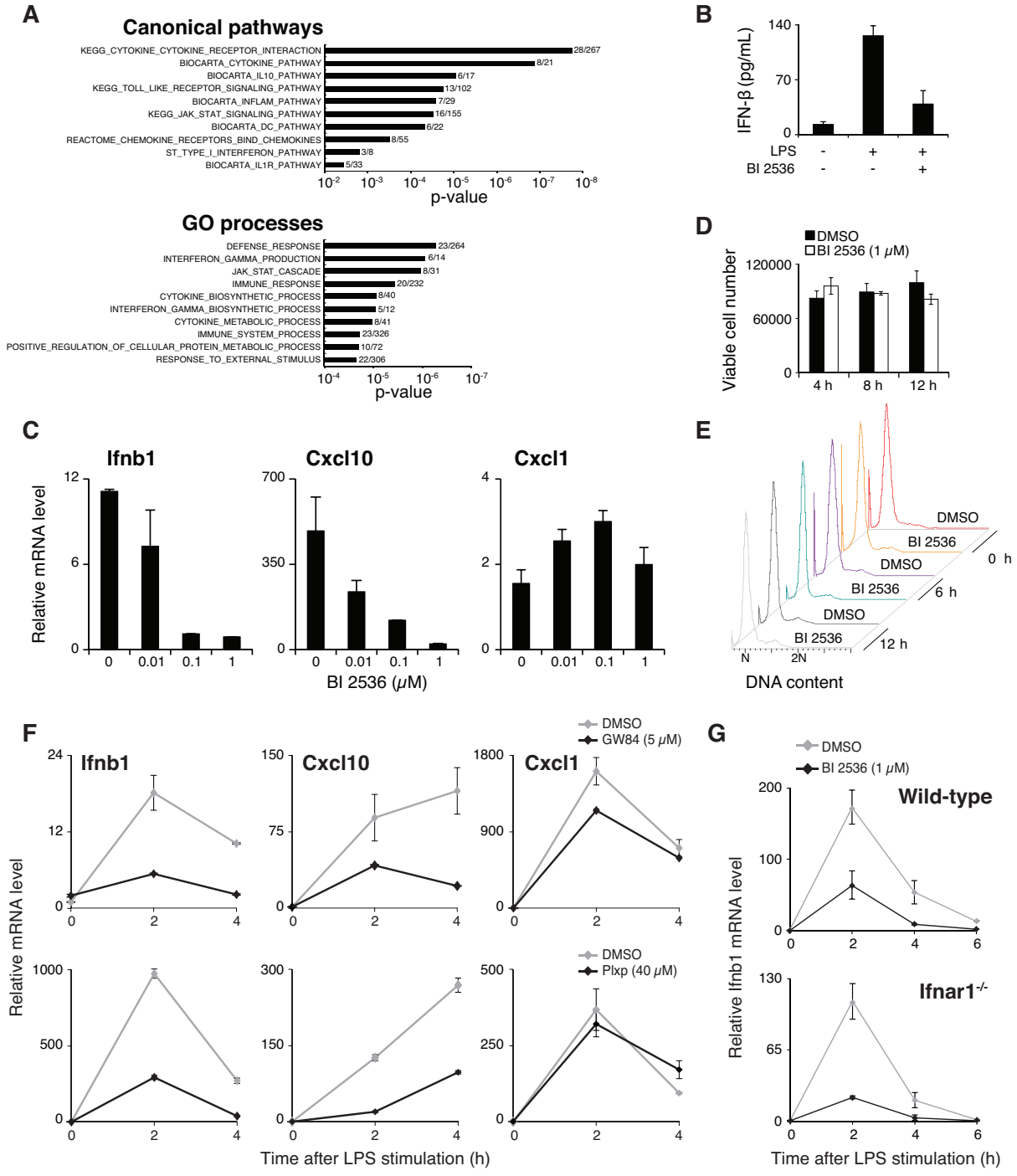

Figure 3.8. BI 2536-Mediated Plk Inhibition Abrogates Antiviral Cytokine Production at the Protein and mRNA Levels, without Affecting the Viability and Cell Cycle Status of DCs

(A) Gene enrichment analysis of BI 2536-dependent genes from microarray measurements. Overlaps between the 311 unique genes downregulated 3-fold by BI 2536 treatment upon LPS or poly(I:C) stimulation, and Gene Ontology (GO) processes and canonical pathways (including the KEGG, REACTOME, and BIOCARTE databases present in the Molecular Signatures Database (MSigDB; see Experimental Procedures). Shown are *P* values (X axis) derived from the overlaps (n/N; top of each bar) between the number of queried genes (n) and genes present in indicated genesets (N).

Figure 3.8 (Continued)
(B) BI 2536 strongly inhibits IFN-β secretion by BMDCs. Shown is IFN-β protein concentration (Y axis; measured by ELISA) in the supernatant of BMDCs treated with DMSO vehicle (-) or BI 2536 (1 μM; +), and stimulated with LPS (+) or left unstimulated (-) for 6 h. Three replicates in each experiment; error bars are the standard error of the mean.
(C) BI 2536 inhibits antiviral cytokine mRNA production in a dose-dependent manner. Shown are mRNA levels (Y axis, qPCR; relative to vehicle control treatment) for two antiviral cytokines (*Ifnb1*, *Cxcl10*) and one inflammatory cytokine (*Cxcl1*) following LPS stimulation in BMDCs pre-treated with increasing amounts of BI 2536 (X axis). Three replicates in each experiment; error bars are the standard error of the mean.
(D) BMDC viability is unaffected by Plk inhibition with BI 2536. Shown are viable cell numbers (Y axis, measured by Alamar blue; relative to a standard curve generated using a range of cell densities) after treatment with BI 2536 (white bars) or DMSO vehicle (black bars) at different time points following addition of BI 2536 (X axis). Three replicates in each experiment; error bars are the standard error of the mean.
(E) The cell cycle state of BMDCs remains unchanged upon Plk inhibition with BI 2536. Shown are DNA contents (flow cytometry) of BMDCs stained with propidium iodide (PI) after treatment with BI 2536 or DMSO vehicle control for 0, 6, and 12 h.
(F) Plk inhibitors structurally unrelated to BI 2536 also abrogate transcription of mRNAs for antiviral cytokines following stimulation with LPS. Shown are mRNA levels (qPCR; relative to t = 0) for *Ifnb1*, *Cxcl10* and *Cxcl1* in BMDCs stimulated with LPS and treated with GW843682X (GW84; top) or Poloxipan (Plxp; bottom) (black line), or with DMSO vehicle (grey line) for 1 hour prior to stimulation. Three replicates for each experiment; error bars are the standard error of the mean.
(G) Plks are directly downstream of TLR engagement. Shown are *Ifnb1* mRNA levels (Y axis, qPCR; relative to t = 0) following LPS stimulation for indicated times (X axis) in wild-type (top) and Ifnar1-/- (bottom) BMDCs treated with BI 2536 (1 μM; black) or vehicle control (DMSO; grey).

Treatment with BI 2536 reduced the mRNA levels of *Cxcl10* and *Ifnb1* (by qPCR) and of secreted IFN-β in a dose-dependent manner (**Figure 3.8B and 3.8C**). Conversely, there was no significant change in the mRNA level of the inflammatory cytokine *Cxcl1* (**Figure 3.8C**). Importantly, BI 2536 treatment prior to stimulation neither impacted the viability nor the cell cycle state of BMDCs (**Figure 3.8D and 3.8E**), suggesting that Plk inhibition does not act through cell cycle effects. In addition, two other pan-Plk inhibitors, structurally unrelated to BI 2536, caused a similar strong decrease in the mRNAs of the antiviral cytokines *Ifnb1* and *Cxcl10*, without affecting *Cxcl1* (**Figure 3.8F and Experimental Procedures**). These results, consistent with our knockdowns, strongly suggest that the effects induced by these three small molecules are due to the inhibition of Plks, and not off-target effects.

To validate that Plks are primary components directly downstream of TLR activation, we measured the effect of BI 2536 on *Ifnb1* induction upon LPS challenge in *Ifnar1*$^{-/-}$ DCs that lack

type I IFN receptor (**Figure 3.8G**). We found a strong inhibitory effect in these cells, consistent with Plks acting directly downstream of TLR signals, and not in an autocrine/paracrine feedback loop mediated by IFN receptor signaling. These observations are also supported by a recent phosphoproteomic study, showing that phosphorylated Plk substrates are enriched after 15 min of LPS stimulation in bone-marrow-derived macrophages (Weintz et al., 2010).

Furthermore, BI 2536 inhibited the nuclear translocation of interferon regulatory factor 3 (IRF3), a key antiviral transcription factor (**Figure 3.9**). We monitored the subcellular localization of IRF3 at the single cell level using confocal microscopy (**Experimental Procedures**). To effectively deliver the drug, we plated DCs on vertical silicon nanowires (Shalek et al., 2010) pre-coated with BI 2536 prior to stimulation. Nanowires coated with the vehicle control did not affect the response of BMDCs to TLR activation (**Figure 3.9A and Figure 3.10A**). DCs treated with BI 2536 showed a dose-dependent decrease in IRF3 nuclear translocation following poly(I:C) or LPS stimulation, whereas the control JNK inhibitor SP 600125 had no effect on IRF3 translocation (**Figure 3.9B and 3.9C, and Figure 3.10B**). On the other hand, BI 2536 did not affect NF-κB p65 localization (**Figure 3.9D and 3.9E**). Notably, IRF3 translocation was also decreased when delivering BI 2536 in solution, but to a lesser extent compared to nanowire-mediated delivery (**Figure 3.10C**). This highlights the utility of highly efficient drug delivery methodologies to induce homogeneous effects in single-cell assays using microscopy. Taken together, these observations place *Plk2* and *4* as critical components upstream of a major antiviral transcription factor.

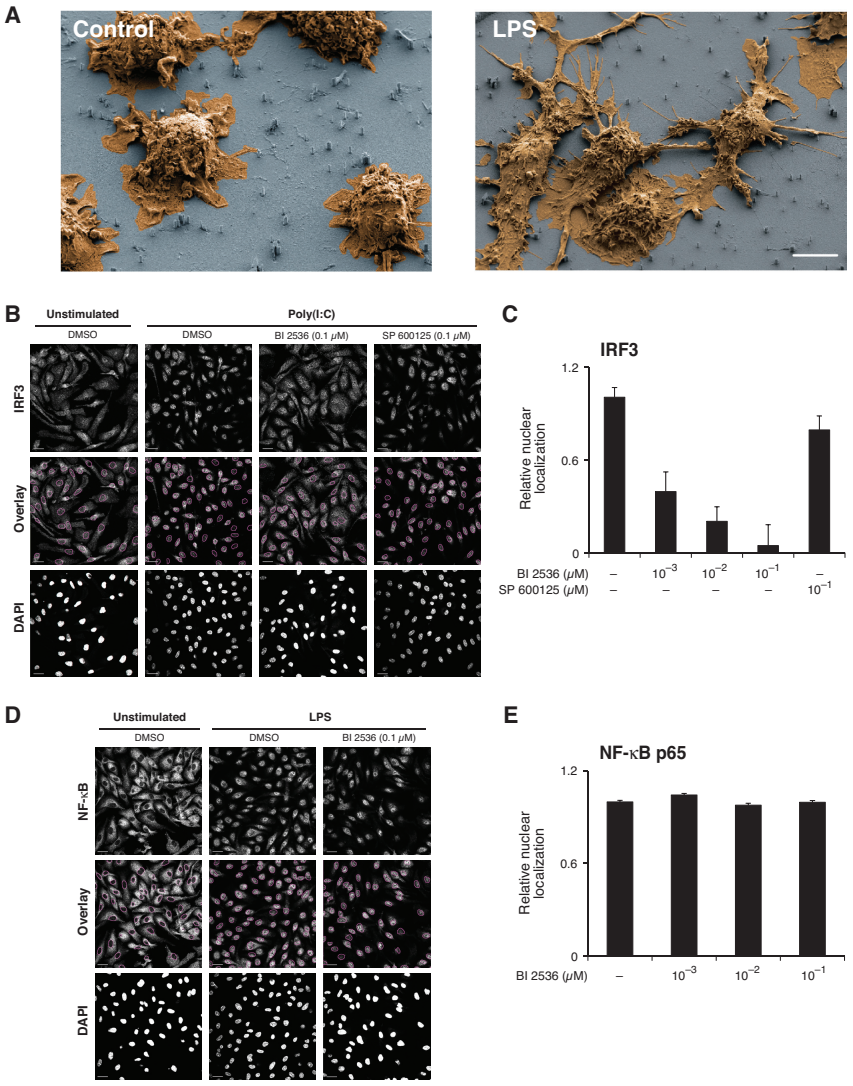

Figure 3.9. BI 2536-Mediated Plk Inhibition Blocks IRF3 Nuclear Translocation in DCs

(A) DCs on nanowires (NW) undergo normal morphological changes upon LPS stimulation. Shown are electron micrographs of BMDCs plated on bare vertical silicon NW that were left unstimulated (left; Control) or stimulated with LPS (right). Scale bars, 5 μm.

(B-E) BI 2536 inhibits IRF3, but not NF-κB p65, nuclear translocation following TLR stimulation. (B and D) Shown are confocal micrographs of BMDCs plated on vertical silicon NW pre-coated with vehicle control (DMSO; B and D), Plk inhibitor (BI 2536; B and D), or control Jnk inhibitor (SP 600125; B), and stimulated with poly(I:C) for 2 h (B) or LPS for 30 min (D) (reflecting peak time of nuclear translocation for IRF3 and NF-κB p65, respectively), or left unstimulated (B and D). Cells were analyzed for DAPI (B and D), IRF3 (B) and NF-κB p65 subunit (D) staining. Scale bars, 5 μM. (C and E) Nuclear translocation (from confocal micrographs) of IRF3 (C) and NF-κB p65 (E) was quantified using DAPI staining as a nuclear mask (purple circles; overlay in B and D) to determine the ratio of total versus nuclear fluorescence (Y axis) in BMDCs cultured on NW coated with different amounts of BI 2536 or SP 600125, or with vehicle control (DMSO; X axis). Three replicates in each experiment; error bars are the standard error of the mean.

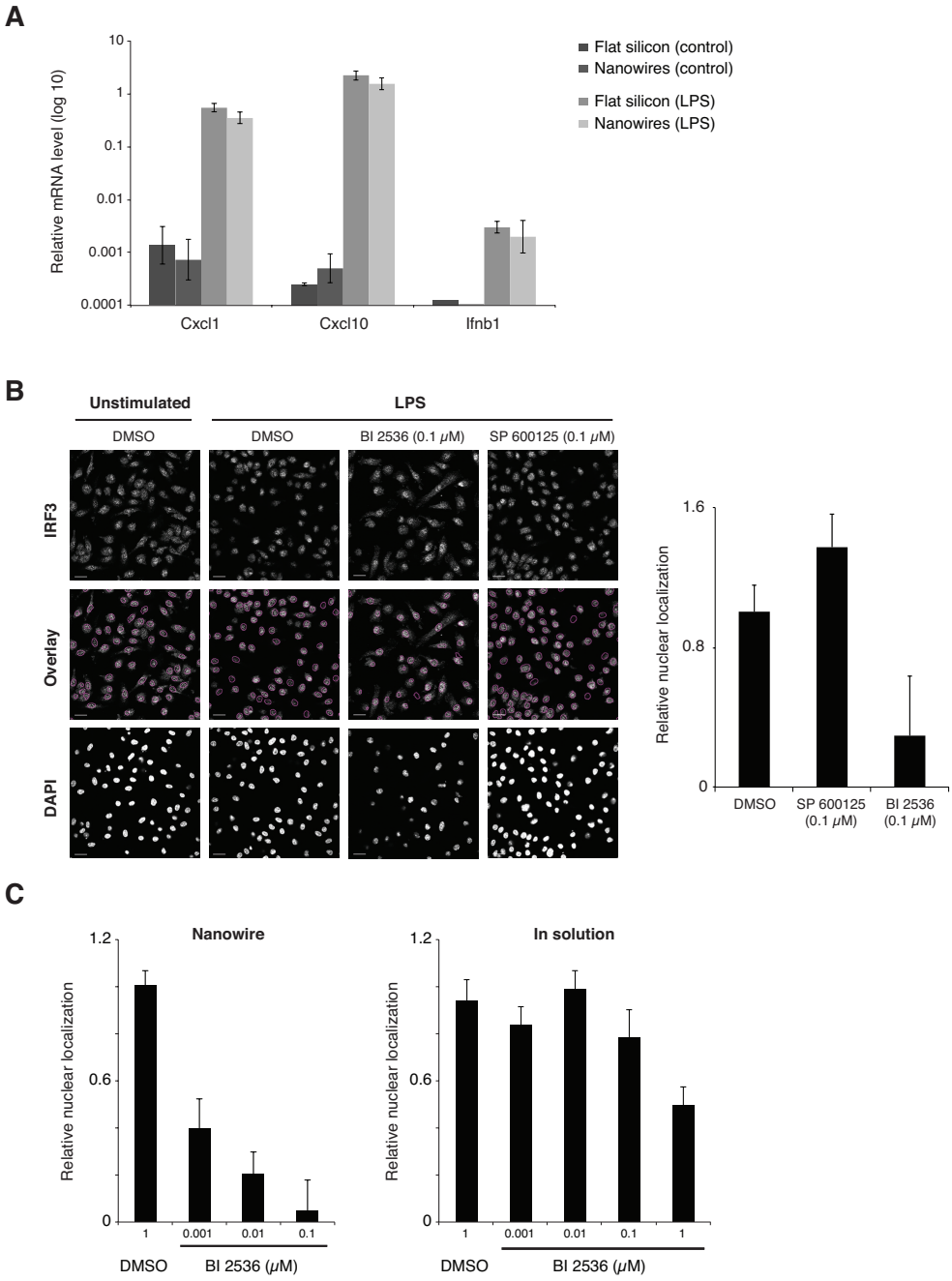

Figure 3.10. BI 2536-Mediated Plk Inhibition Blocks IRF3 Nuclear Translocation in LPS-Stimulated DCs

(A) DCs plated on vertical silicon nanowires (NW) respond normally to TLR stimulation. Shown are cytokine mRNA levels (qPCR; relative to Gapdh mRNA) in BMDCs plated on NW or a flat silicon surface, and stimulated (LPS) or left untreated (control). Left to right: *Cxcl1*, *Cxcl10*, *Ifnb1*. Three replicates in each experiment; error bars are the standard error of the mean.

(B) BI 2536 inhibits IRF3 nuclear translocation following LPS stimulation. Shown are confocal micrographs (left panel) of BMDCs plated on vertical silicon NW pre-coated with vehicle control (DMSO), Plk inhibitor (BI 2536), or control Jnk inhibitor (SP 600125), and stimulated with LPS for 45 min (reflecting peak time of nuclear translocation for IRF3 in the context of LPS stimulation), or left unstimulated. Cells were analyzed for DAPI and IRF3 staining. Scale bars, 5 μM. Nuclear translocation (from confocal micrographs) of IRF3 was quantified (right panel) using DAPI staining as a nuclear mask (purple circles

Figure 3.10 (Continued)
on micrographs) to determine the ratio of total versus nuclear fluorescence (Y axis) in BMDCs cultured on NW coated with BI 2536, SP 600125, or vehicle control (DMSO; X axis). Three replicates in each experiment; error bars are the standard error of the mean.
(C) Decrease in IRF3 nuclear translocation may be more efficient with NW-mediated delivery of BI 2536 than with delivery in solution. Shown are quantifications of confocal micrographs from BMDCs plated on vertical NW pre-coated with different amounts of BI 2536 (Nanowire; left panel) or left blank to allow in-solution delivery of BI 2536 (In solution; right panel). Cells were stimulated with poly(I:C) for 2 h prior to staining for DAPI and IRF3 as in B.

3.8. Plks Are Essential for Activation of All Well-Established IFN-Inducing Pathways in Conventional and Plasmacytoid DCs

DCs can be broadly categorized into two major subtypes – conventional and plasmacytoid DCs – each relying on distinct mechanisms to induce type I IFNs and antiviral gene expression (Blasius and Beutler, 2010; Luber et al., 2010). In conventional DCs (cDCs), antiviral responses are activated through TLR4/3 signaling (via TRIF), or through the cytosolic sensors RIG-I or MDA-5 (via MAVS) (**Figure 3.11A**). In plasmacytoid DCs (pDCs; specialized IFN-producing cells), the antiviral response depends solely on endosomal TLR7 and 9 that signal via MYD88 (**Figure 3.11A**) (Blasius and Beutler, 2010; Kato et al., 2005; Kato et al., 2006; Takeuchi and Akira, 2010).

BI 2536 treatment showed that Plks are essential for the viral-sensing pathways in both cDCs and pDCs. In cDCs, BI 2536 inhibited the transcription of antiviral genes (*Ifnb1* and *Cxcl10*) upon infection with each of four different viruses: vesicular stomatitis virus (VSV) (**Figure 3.11B, top**), Sendai virus (SeV; **Figure 3.12A top**), or Newcastle disease virus (NDV; **Figure 3.12A bottom**), all three sensed through RIG-I, and encephalomyocarditis virus (EMCV), sensed through MDA-5 (**Figure 3.11B, bottom**).

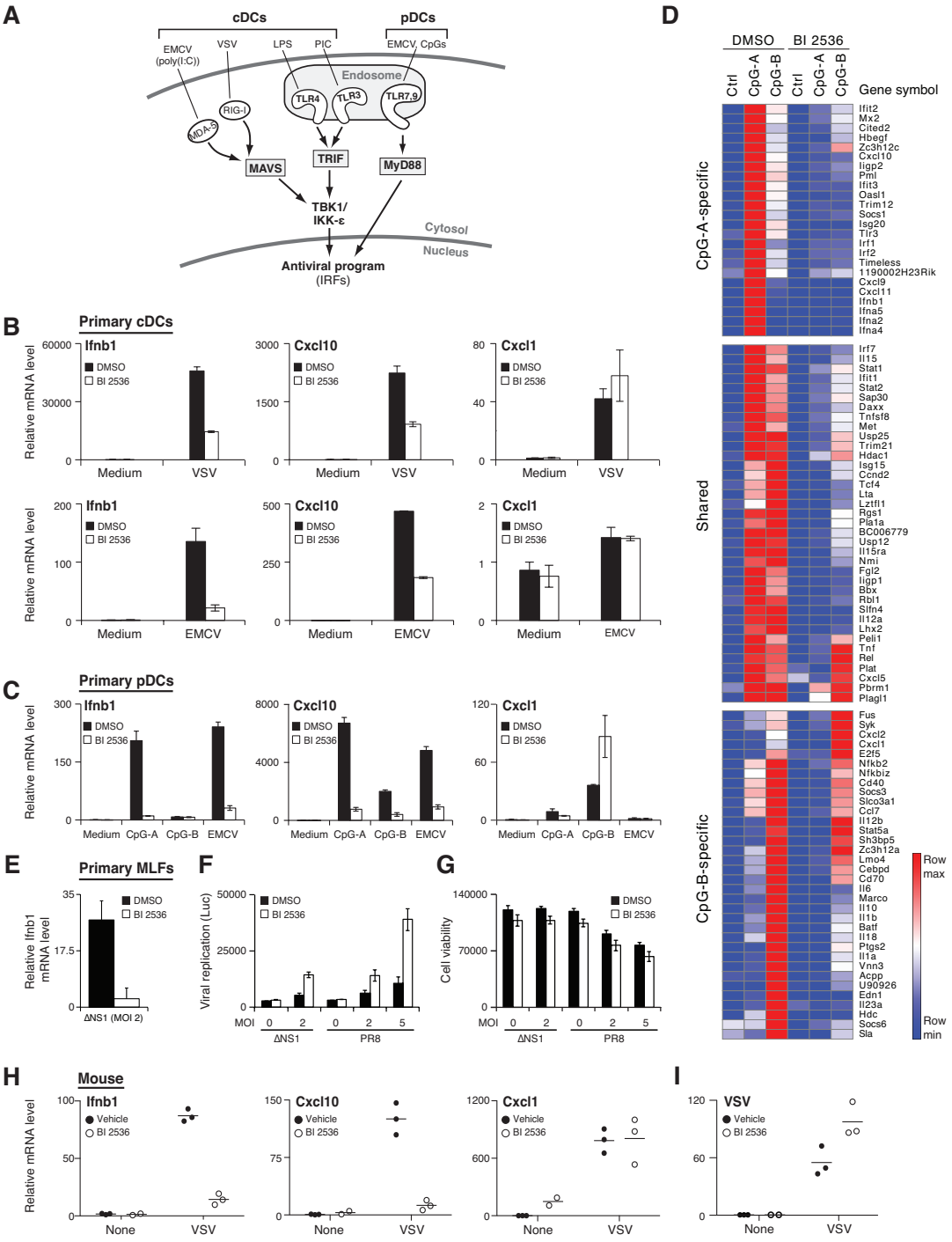

Figure 3.11. PIks Are Critical in the Induction of Type I Interferons *In Vitro* and *In Vivo*

(A) IFN-inducing pathways in conventional DCs (cDCs) and plasmacytoid DCs (pDCs) (Blasius and Beutler, 2010).

(B, C) BI 2536 inhibits mRNA levels for antiviral cytokines in response to diverse stimuli in cDCs and pDCs. Shown are *Ifnb1*, *Cxcl10* and *Cxcl1* mRNA levels (qPCR; relative to t = 0) in cells treated with BI 2536 (1 μM; white bars) or DMSO vehicle (black bars) in cDCs (B) infected with VSV (MOI 1; B top) or with EMCV (MOI 10; B bottom), and in pDCs (C) stimulated with CpG type A or B, or infected with EMCV (MOI 10). Three replicates in each experiment; error bars are the standard error of the mean.

Figure 3.11 (Continued)
(D) BI 2536 inhibits the CpG-A response, but has little effect on the CpG-B response. Shown are mRNA levels (nCounter) for the 118 TLR signature genes (rows) in pDCs treated with DMSO vehicle or BI 2536 (1 μM) and left untreated (Ctrl) or stimulated with CpG-A or -B for 6 h (columns). Three clusters of genes are shown: CpG-A-specific (top), CpG-B-specific (bottom), and shared by CpG-A and -B (middle).
(E-G) BI 2536 inhibits IFN-β production in primary mouse lung fibroblasts (MLFs), leading to an increase in viral replication. MLFs treated with BI 2536 (1 μM; white bars) or vehicle control (DMSO; black bars) were infected with influenza ΔNS1 or PR8 strains at indicated MOIs. Shown are Ifnb1 mRNA levels measured by qPCR (relative to t = 0; E), viral replication as measured by luciferase (Luc) activity in reporter cells (F), and cell viability measured by CellTiter-Glo assay (G).
(H and I) BI 2536 inhibits antiviral cytokine mRNA production, while increasing viral replication during *in vivo* VSV infection. Shown are *Ifnb1*, *Cxcl10* and *Cxcl1* mRNA (H), and VSV viral RNA (I) levels (qPCR; relative to uninfected animals) from popliteal lymph nodes of mice injected with BI 2536 (white circles) or DMSO vehicle (black circles) prior to and during the course of infection with VSV (intra-footpad). Nodes were harvested six hours post-infection. Each circle represents one animal (n = 3). Data is representative of three independent experiments for each condition.

Notably, BI 2536 neither affected the mRNA level of *Cxcl1* (an inflammatory cytokine) in response to any of these four viruses, nor affected the response to heat-killed *Listeria monocytogenes*, a natural TLR2 agonist (**Figure 3.11B and Figure 3.12A and 3.12B**). In pDCs, BI 2536 treatment nearly abrogated the transcription of mRNAs for the antiviral cytokines *Ifnb1*, *Ifna2*, and *Cxcl10* after stimulation with type A CpG oligonucleotides (CpG-A), or infection with EMCV, sensed by TLR9 and 7, respectively (**Figure 3.11C, Figure 3.12C**). Conversely, in pDCs stimulated with CpG-B – a ligand known to activate inflammatory pathways but not IFN-inducing pathways – BI 2536 treatment decreased *Cxcl10* mRNA while moderately increasing *Cxcl1* mRNA (**Figure 3.11C**). Finally, we measured the impact of BI 2536 treatment on the 118-gene signature in pDCs. BI 2536 strongly inhibited genes that are typically induced by CpG-A alone or by both CpG-A and -B, while having a minor effect, if any at all, on CpG-B-specific genes in pDCs (**Figure 3.11D**). Our findings thus contribute to revealing the molecular determinants of IFN production in pDCs, which are still poorly characterized (Reizis et al., 2011). More importantly, these results demonstrate a critical role for Plks across all well-known IFN-inducing pathways.

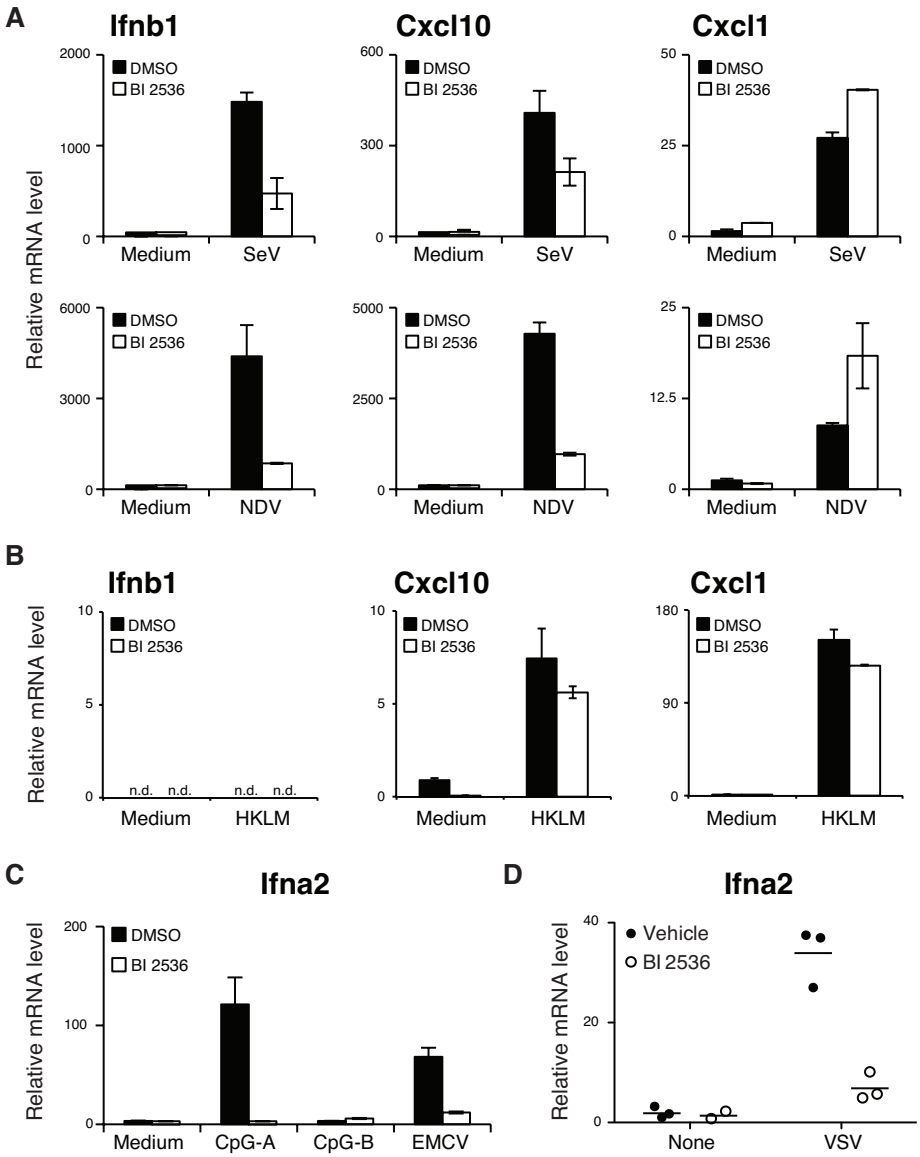

Figure 3.12. Plks Are Critical in Antiviral Responses *In Vitro* and *In Vivo*

(A) Plks are critical in RIG-I-mediated antiviral responses *in vitro* in DCs. Shown are mRNA levels (qPCR; relative to control, "medium") in conventional DCs (GM-CSF-induced BMDCs) treated with BI 2536 (white bars) or DMSO vehicle (black bars), and infected at a multiplicity of infection (MOI) 1 with Sendai virus (SeV; top) or Newcastle disease virus (NDV; bottom). Three replicates in each experiments; error bars are the standard error of the mean.

(B) Plk inhibition does not affect DC response to *Listeria monocytogenes*, a natural TLR2 agonist. Shown are mRNA levels (qPCR; relative to t = 0) for *Ifnb1*, *Cxcl10* and *Cxcl1* in BMDCs stimulated with heat-killed *Listeria monocytogenes* (HKLM; MOI 5) and treated with BI 2536 (white bars), or with DMSO vehicle (black bars) for 1 hour prior to stimulation. Three replicates for each experiment; error bars are the standard error of the mean.

(C) Plks are critical in type I interferon α2 (*Ifna2*) gene production by plasmacytoid DCs (pDCs). Shown is the mRNA level (qPCR; relative to control, "medium") of Ifna2 in pDCs (Flt3L-induced BMDCs) treated with BI 2536 (1 μM; white bars) or DMSO control (black bars), and stimulated with CpG-A or -B, or infected with EMCV (MOI 10). Three replicates in each experiment; error bars are the standard error of the mean.

Figure 3.12 (Continued)
(D) Plk inhibition *in vivo* inhibits type I IFN α2 production in the lymph node. Shown is *Ifna2* mRNA level (qPCR; relative to uninfected animals) from popliteal lymph nodes of mice injected with BI 2536 (white circles) or DMSO vehicle (black circles) prior to and during the course of infection with VSV intra-footpad. Nodes were harvested six hours post-infection. Each circle represents one animal (n = 3). Data is representative of two or three independent experiments for each condition.

3.9. Plks Are Essential in the Control of Host Antiviral Responses

To evaluate the impact of Plk inhibition on the outcome of viral infection, we infected primary mouse lung fibroblasts (MLFs) with influenza virus (**Experimental Procedures**). In the presence of BI 2536 treatment, primary MLFs infected with influenza failed to produce interferon (**Figure 3.11E**), and showed elevated replication of both a wild type and a typically poorly replicating mutant virus (**Figure 3.11F**). The reduced interferon response was not due to drug-induced toxicity (**Figure 3.11G**).

Next, we tested the effects of Plk inhibition during viral infection in mice *in vivo*. BI 2536 exhibits good tolerability in mice (Steegmaier et al., 2007) and humans (Mross et al., 2008), and is currently in Phase II clinical trials as an anti-tumor agent in pancreatic cancer, hormone-refractory prostate cancer, and non-small cell lung cancer (Strebhardt, 2010). Given this efficacy and safety of BI 2536 *in vivo*, we tested whether Plk inhibition would also affect the response to viral infection in animals. In mice infected with VSV (**Experimental Procedures**), injections of BI 2536 strongly suppressed mRNA production in popliteal lymph nodes for type I IFNs (*Ifnb1*, *Ifna2*) and *Cxcl10*, whereas *Cxcl1* mRNA induction remained unchanged (all compared to vehicle control; **Figure 3.11H and Figure 3.12D**). Concomitantly, VSV replication in the lymph node rapidly increased as reflected by elevated VSV RNA levels (**Figure 3.11I**). This increase in VSV RNA in the presence of BI 2536 is comparable to the observed phenotype of VSV-infected *Ifnar1*$^{-/-}$ animals (Iannacone et al., 2010). Because in the VSV model used here type I IFNs are produced by both infected CD169^{+} subcapsular sinus macrophages and pDCs (Iannacone et

al., 2010), we cannot distinguish whether Plk inhibition affects macrophages, pDCs, or both. Nevertheless, our results confirm the physiological importance of Plks in the host antiviral response in both *ex vivo* primary MLFs and *in vivo* mouse lymph nodes.

3.10. Plks Affect the Phosphorylation of Dozens of Proteins Post-LPS Stimulation, Including Known and Candidate Antiviral Regulators

We next sought to discover the signaling pathways between Plks and antiviral gene transcription. Notably, the only signaling molecule previously reported as shared across well-known IFN-inducing pathways is TRAF3 (Oganesyan et al., 2006), which links TLR or RLR adaptors (*i.e.*, MYD88, TRIF, MAVS) to downstream kinase cascades and the transcription of type I IFNs (Blasius and Beutler, 2010; Takeuchi and Akira, 2010). We used MicroWestern Arrays (MWAs) (Ciaccio et al., 2010) to measure changes in the phosphorylation levels of 20 signaling phosphoproteins (using 23 phosphosite-specific antibodies) of the TLR pathway (**Experimental Procedures**). We also measured total protein levels for 6 proteins. We measured phosphorylation and protein levels in BMDCs at each of 12 combinations of four time points (0, 20, 40, 80 min after LPS stimulation) and three perturbations (vehicle control, BI 2536, and SP 600125, a JNK inhibitor serving as a negative control). Although, LPS stimulation alone led to the expected changes in phosphorylation states and protein levels, compared to pre-stimulation levels (*e.g.*, early peak of phosphorylation for ERK1/2, p38, and Mapkapk2, and rapid degradation of IκBα; **Figure 3.13A**), BI 2536 treatment did not cause any significant changes in the phosphorylation levels of any of these 20 proteins, compared to vehicle or to the negative control JNK inhibitor SP 600125 (**Figure 3.13A and Figure 3.14A and 3.14B**). This may reflect the paucity of known phospho-proteins and phospho-sites in these pathways, and consequently the limited number of available phospho-specific antibodies targeting known TLR signaling nodes. We therefore hypothesized that Plks could affect previously unrecognized

regulators of IFN-inducing pathways and/or known regulators with no existing antibodies to specific phosphosites.

To test this hypothesis and overcome the limited availability of phospho-specific antibodies, we used SILAC-based unbiased phosphoproteomics (Villen and Gygi, 2008) to identify any proteins whose phosphorylation depends on Plks in response to TLR stimulation (**Figure 3.13B top**). We compared the levels of phospho-tyrosine, -threonine and -serine peptides following stimulation with LPS (for 30 or 120 min) in BMDCs pre-treated with BI 2536 versus those treated with vehicle (DMSO). We identified and quantified 5,061 and 5,997 phosphopeptides after 30 and 120 minutes, respectively, for a total of 10,236 individual phosphosites (**Figure 3.13B**). BI 2536 treatment led to a significant ($P < 0.001$) increase or decrease in the level of 510 phosphopeptides, reflecting 413 distinct proteins (**Figure 3.13B**). Thus, BI 2536-mediated inhibition of Plks had a substantial effect on the TLR phosphoproteome of DCs.

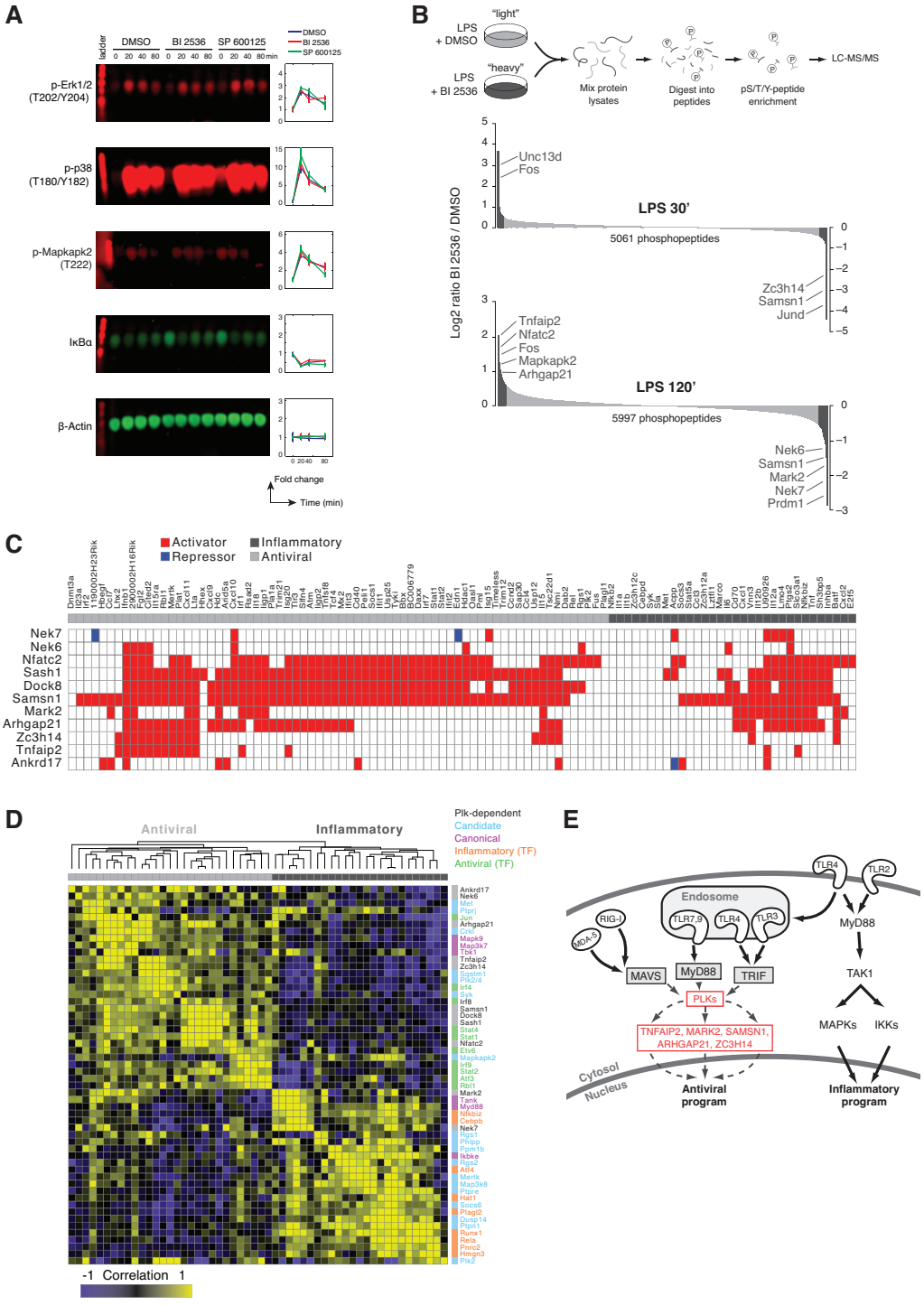

Figure 3.13. Unbiased Phosphoproteomics Identifies a Plk-Dependent Antiviral Pathway

(A) BI 2536 does not affect phosphorylation and protein levels of known TLR signaling nodes. Shown are representative MicroWestern Array (MWA; see Experimental Procedures) blots (left) obtained from analyzing lysates from BMDCs pre-treated with DMSO, BI 2536 (1 μM), or SP 600125 (5 μM) and stimulated with LPS for 0, 20, 40, 80 min. Blots were analyzed using indicated antibodies (left most), and fold change in fluorescence signals was quantified relative to t = 0 (right; see Experimental Procedures). Error bars are the standard error of the mean of triplicate MWA blots.

Figure 3.13 (Continued)
(B) BI 2536 affects protein phosphorylation levels during LPS stimulation. Top: Schematic depiction of experimental workflow. From left to right: LPS-stimulated BMDCs cultured in "heavy" or "light" SILAC medium were pre-treated with BI 2536 (1 μM) or DMSO, respectively. Protein lysates were mixed (1:1) and digested into peptides with trypsin, before phospho-serine, -threonine and -tyrosine (pS/T/Y) peptide enrichment, and LC-MS/MS analysis. Bottom: Shown are the differential phosphorylation levels (average $\log_2$ ratios of two independent experiments; Y axis) of all 5061 and 5997 phosphopeptides respectively identified and quantified by LC-MS/MS (X axis) at 30 min (top) and 120 min (bottom) post-LPS stimulation. Dark grey: phosphopeptides with a significant change ($P_{unadjusted} < 0.001$ for both time points; $FDR_{30min} = 0.05$; $FDR_{120min} = 0.03$; left: induced; right: repressed). Average ratios from phosphopeptides identified and quantified in two independent experiments are depicted.
(C) Eleven Plk-dependent phosphoproteins significantly affect the expression of TLR signature genes. Shown are significant changes in expression of the TLR signature genes (rows) following knockdown of each of the 11 phosphoproteins (columns), following stimulation with LPS for 6 h. Red and blue mark significant hits (as presented in **Figure 3.3**) and are shown only for genes where the effect was consistent between two independent experiments.
(D) Functional characterization based on similarity of perturbation profiles. Shown is a correlation matrix of the perturbation profiles from C (grey), and those from Figure 2B including canonical (purple) and candidate (blue) signaling components as well as core antiviral (green) and inflammatory (orange) transcriptional regulators. Yellow: positive correlation; purple: negative correlation; black: no correlation.
(E) A Plk-dependent pathway in antiviral sensing. Shown is a diagram of a model of the Plk-dependent pathway of IFN induction in innate immunity. Out of the 11 Plk-dependent proteins described in C and D, only the 5 showing a phenotype with 2 independent shRNAs are depicted.

Of note, 35% (2489/7018) of the phospho-sites we identified were recently reported in mouse bone marrow-derived macrophages treated with LPS (**Figure 3.14C, left**) (Weintz et al., 2010), and 483 of our phosphosites were among 1858 sites (26%) reported in a phosphoproteomic study of LPS signaling in a macrophage cell line (**Figure 3.14C, left**) (Sharma et al., 2010). A comparison of the phosphosites of known kinases showed similar overlaps between the three studies (**Figure 3.14C, right**). This level of agreement across these two studies, despite differences in cell types, timepoints, and experimental procedures, supports our results.

The Plk-dependent phosphoproteins include several known regulators of antiviral pathways (*e.g.*, *Prdm1*, *Fos*, *Unc13d*) (Crozat et al., 2007; Keller and Maniatis, 1991; Takayanagi et al., 2002), as well as many additional protein candidates with no previously known function in viral-sensing (**Figure 3.13B**). Notably, proteins involved in the TBK1/IKK-ε/IRF3 axis were detected and quantified but did not show any significant change in phosphorylation levels upon Plk inhibiton, consistent with the MicroWestern array data. Conversely, Plk inhibition with BI 2536

decreased the phosphorylation levels of cell cycle regulators of the Jun family of transcriptional regulators (*i.e.*, *Jund*) that we recently reported as co-opted by antiviral pathways (Amit et al., 2009). BI 2536 treatment also decreased the phosphorylation levels of the mitotic kinases *Nek6* and *Nek7* (**Figure 3.13B**). The recent observation that the phosphorylation *Nek6* substrates are increased following LPS stimulation in macrophages indirectly corroborates our finding that *Nek6* may be active in the TLR signal flow (Weintz et al., 2010). To test the role of these new candidates, we returned to our shRNA perturbation-based approach.

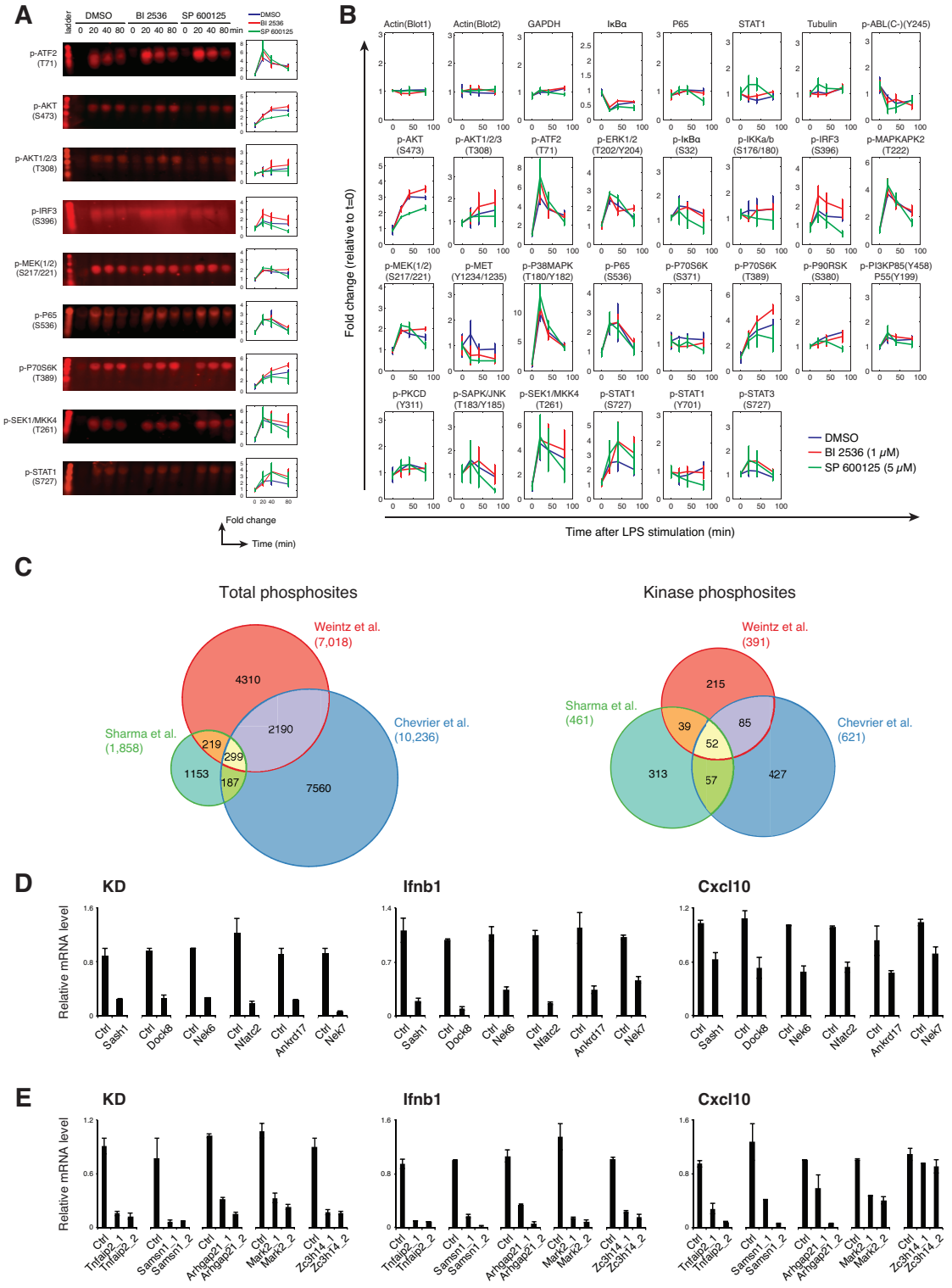

Figure 3.14. Plk Inhibition Does Not Affect Known TLR Signaling Components but Affects Eleven Newly Identified Plk-Dependent Phosphoproteins

(A, B) BI 2536-mediated Plk inhibition does not affect protein and/or phosphorylation levels of known TLR signaling nodes. (A) Shown are representative MicroWestern Array (MWA; see Experimental Procedures) blots obtained from analyzing lysates from BMDCs pre-treated with DMSO, BI 2536 (1 μM), or SP 600125

Figure 3.14 (Continued)
(5 μM) and stimulated with LPS for 0, 20, 40, 80 min. Blots were analyzed using indicated antibodies (left most), and fold change in fluorescence signals was quantified relative to t = 0 (right; see Experimental Procedures). Error bars are the standard error of the mean of triplicate MWA blots. (B) Shown are the differential protein and phosphorylation levels (fold change; Y axis) of 6 proteins and 23 phosphosites in BMDCs treated with BI 2536 (red line), control JNK inhibitor (SP 600125; green line), or DMSO vehicle (blue line), and stimulated with LPS (0, 20, 40, 80 min; X axis). Band intensities on MWA blots were quantified using Li-cor Odyssey analysis software (Experimental Procedures). For each antibody, data was normalized to β-actin levels; error bars are the standard error of the mean of triplicate MWA blots.
(C, D) 11 Plk-dependent phosphoproteins are critical for TLR-mediated antiviral responses in DCs. Shown are mRNA levels (qPCR; relative to non-targeting control shRNAs, Ctrl) for knockdown (KD) efficiency (left), Ifnb1 (middle), and Cxcl10 (right) in BMDCs following LPS stimulation. Genes with one and two shRNAs are shown in C and D, respectively. Three replicates in each experiment; error bars are the standard error of the mean.
(E) Comparison of phosphosites identified in our study and in two recent reports (Weintz et al., and Sharma et al.). Shown are proportional Venn diagrams of the total unique phosphosites identified by the 3 studies (left), and the phosphosites harbored by kinases only (right). Total numbers of unique phosphosites per study are indicated in parentheses.

3.11. Plk-Dependent Phosphoproteins Affect the Antiviral Response

We tested 25 novel Plk-dependent phosphoproteins, using shRNA perturbation in BMDCs followed by both qPCR and the 118 TLR gene signature measurements. These 25 candidates satisfied three criteria: (**1**) there was no prior knowledge of their function in viral-sensing pathways; (**2**) their phosphoprotein levels were consistently up- or down-regulated upon BI 2536 treatment (in two independent experiments); and (**3**) they had detectable mRNA expression and/or differential expression upon stimulation.

Of the 18 phosphoproteins showing efficient knockdown, 11 caused a significant decrease in *Ifnb1* mRNA levels with a single shRNA (*Sash1*, *Dock8*, *Nek6*, *Nek7*, *Nfatc2*, and *Ankrd17*; **Figure 3.14D**), or with two independent shRNAs (*Tnfaip2*, *Samsn1*, *Arhgap21*, *Mark2*, and *Zc3h14*; **Figure 3.14E**). Decrease in *Cxcl10* expression was less prominent, consistent with our previous observations of BI2536's weaker effect on this cytokine during LPS stimulation (**Figure 3.14D and 3.14E, far right panels**). Each of the 11 Plk-dependent phosphoproteins tested affected at least 9 targets in the 118-gene signature (on average, 39 targets ± 30 SD; **Figure 3.13C**), and 9 affected more than 10% of the targets in the TLR gene signature (**Figure 3.13C**).

9 of the 11 Plk-dependent phosphoproteins affected the TLR signature comparably to major antiviral regulators (**Figure 3.13D**). For example, the profiles of newly identified *Samsn1*, *Dock8*, and *Sash1* were closely correlated to those of *Stat* and *Irf* family members (**Figure 3.13D**). The profile of the *Plk2/4* double knockdown was most correlated to those of *Tnfaip2* and *Zc3h14*, proteins without a known molecular function. Interestingly, *Tnfaip2* has been associated with rheumatoid arthritis and autoimmune myocarditis in recent genome-wide association studies (Consortium, 2007; Kuan et al., 1999). Our findings provide a potential molecular context for this disease association.

Taken together, these results establish a previously unrecognized role for Plks in pathways inducing type I IFNs *ex vivo* in several primary cell types (*i.e.*, DCs, MLFs), and *in vivo* in a mouse VSV infection model, likely by controlling the phosphorylation and activity of a new module of at least 11 components (**Figure 3.13E**).

3.12. Discussion

Therapeutic targeting of TLR pathways to dampen or enhance immune responses is an area of intense investigation in infectious disease, vaccine development, and autoimmune disorders (Hennessy et al., 2010; Kanzler et al., 2007; Ulevitch, 2004). Using an integrative strategy that combines transcriptional profiling, genetic and small molecule perturbations, and phosphoproteomics, we discovered 19 new signaling regulators not previously implicated in the TLR network (including 8 based on mRNA changes – 7 of 17 initial candidates together with Plk4; and 11 based on phosphorylation changes). Among these new regulators, we further explored the roles of Crkl and Plks and uncovered their critical roles as antiviral regulators. In particular, using the Plk inhibitor BI 2536, we demonstrated a significant role for Plks in host

defense against viral infections *in vitro* and *in vivo* in mice. We showed that Plks are essential for all well-described IFN-inducing pathways (*i.e.*, TLR- and RLR-mediated), and control a previously uncharacterized module of at least eleven members, including a functionally uncharacterized gene (*i.e.*, *Tnfaip2*), that has been genetically associated with two autoimmune diseases (**Figure 3.15**).

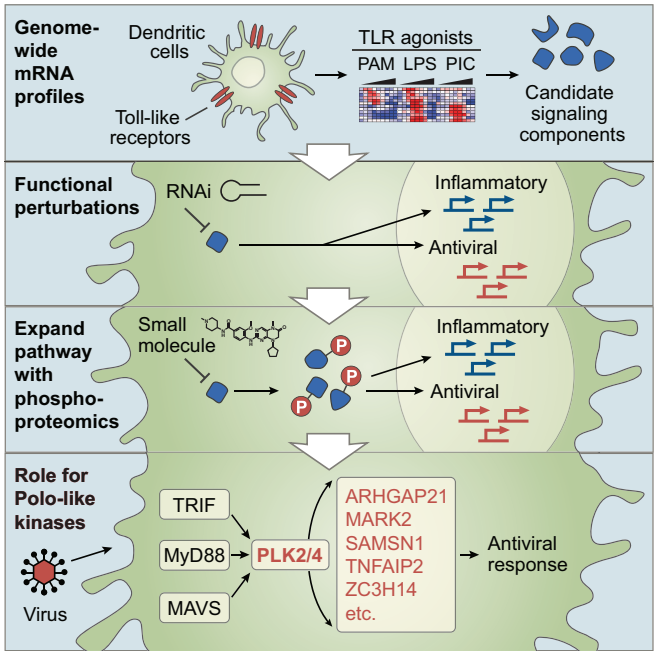

Figure 3.15. Graphical Summary of the Approach and Main Results from Chapter 3

Crkl is an SH2- and SH3-adaptor protein from the Crk family of phosphotyrosine adaptors, and plays a known role in cell proliferation, cell adhesion, and regulation of gene expression (Birge et al., 2009). Crk adaptors have been linked with immune signaling in both normal (*e.g.*, lymphocyte activation, IFN signaling) and pathological (*e.g.*, diabetes) conditions (Birge et al., 2009; Laloraya et al., 2006). We report a new role for Crkl in modulating the antiviral arm of TLR4 signaling. This may explain recent findings that Crkl is targeted by the non-structural protein 1 (NS1) of influenza A virus (Heikkinen et al., 2008; Hrincius et al., 2010). We also find that in response to DC stimulation, Crkl is phosphorylated on Tyrosine 132, a known

phosphorylation site in other responses (Ballif et al., 2008). Future studies are needed to characterize the role of this phosphorylation site in Crkl-dependent regulation of antiviral responses, to determine its molecular targets, and to find the upstream tyrosine kinase. Both the similarity of perturbational profiles and previous observations (Birge et al., 2009; Hrincius et al., 2010) suggest that Jnk2 is a likely candidate partner for Crkl in the TLR system.

We find that Plk2 and 4 – typically associated with cell cycle regulation and proliferation – have been co-opted in differentiated DCs to serve as major regulators of antiviral pathways. This is consistent with our earlier finding of the role of cell cycle transcription factors in the antiviral responses (Amit et al., 2009). In addition, we find that two other cell cycle kinases, Nek6 and Nek7, regulate IFN transcription. Given the 85% homology between the kinase domain of Nek6 and Nek7, they may have partially redundant roles in regulating TLR pathway substrates, and hence, combinatorial perturbation of these two kinases might reveal different and critical aspects of their roles in the TLR system.

Despite substantial progress in elucidating the role of Plk1 in the regulation of mitotic events, the functions of its paralogs – Plk2, 3, and 4 – remain poorly defined (Strebhardt, 2010). Plk2, 3, and 4 participate in cell cycle processes to a lesser extent than Plk1, and their new roles in non-dividing cells have begun to emerge, especially in the nervous system (Archambault and Glover, 2009; Seeburg et al., 2005). In immune DCs, we show that Plk2 and 4 have redundant functions in the control of antiviral responses, illustrating the importance of epistasis analysis of mammalian signaling networks. Such redundancies might be obscuring additional roles for Plks, Neks, and other kinase families across multiple systems. However, genetic interactions currently cannot be found by genome-wide screens in mammalian cells. Thus, a more systematic computational and experimental approach is required in order to predict and test them.

BI 2536 blocked the nuclear translocation of the IRF3 transcription factor without affecting its phosphorylation level (based on MicroWestern arrays and phosphoproteomics). Similar observations have very recently been reported in innate immune pathways. Indeed, the sodium-potassium ATPase inhibitor bufalin inhibits the nuclear translocation of NF-κB upon TNF stimulation without affecting its phosphorylation level (Ye et al., 2011). Our results with BI 2536 thus demonstrate that IRF3 translocation is likely to be regulated by a mechanism that does not impact phosphorylation.

One of the goals of our work was to identify signaling nodes with clear regulatory functions that could be potential therapeutic targets. We find that Plk inhibition suppresses type I IFN production *in vivo* during viral infection, a finding that has potential clinical implications. For example, IFN receptor knockouts have reduced symptoms in a mouse model of Systemic Lupus Erythematosus (SLE) (Santiago-Raber et al., 2003), and disease activity in patients correlates with IFN expression signatures (Banchereau and Pascual, 2006). Furthermore, TLR7 and 9 expressed by pDCs and B cells are well-known to be involved in SLE pathogenesis, and a bifunctional inhibitor (IRS 954) of both TLR7 and 9 increases the survival of lupus-prone animals (Barrat and Coffman, 2008). Our *in vitro* experiments demonstrate that Plk inhibition strongly suppresses both TLR7 and 9 responses in pDCs. Thus, testing the effect of BI 2536 on a mouse model of lupus, such as (NZB×NZW)F1 mice, will be critical in assessing the potential therapeutic implications of Plk inhibition for SLE. Inhibiting Plks might also be relevant to other disorders with TLR7- and TLR9-driven components such as psoriasis (Lande et al., 2007) and AIDS (Mandl et al., 2008).

Our study demonstrates how an integrative strategy – combining transcriptomics, genetic and chemical perturbations, and unbiased phosphoproteomics – can be used to discover and place new components of signaling networks. This approach may also be broadly applicable for

characterizing the functions of genes reported in genome-wide association studies (*e.g.*, *Tnfaip2*), for uncovering new potential therapeutic targets (*e.g.*, Plks), and for re-purposing existing small molecules in new physiological contexts (*e.g.*, using the cancer drug BI 2536 to repress innate immune responses). The vast public compendia of microarray data could serve as starting points for identification of relevant signaling components in diverse biological systems, followed by perturbations and signature measurements. Nevertheless, since the mRNAs corresponding to many pathway components do not change upon pathway activation, our approach is far from exhaustive. Combination of our perturbation-based approach with large-scale biochemical measurements (*e.g.*, post-translational modifications, protein-protein interactions), will lead to more comprehensive, integrated maps of signaling and transcriptional networks.

3.13. Experimental Procedures

Preparation of dendritic cells

Bone marrow-derived dendritic cells (BMDCs) were generated from 6-8 week old female C57BL/6J mice (Jackson Laboratories). Bone marrow cells were collected from femora and tibiae and plated at 10^6 cells/mL on non-tissue culture treated petri dishes in RPMI-1640 medium (Gibco), supplemented with 10% FBS, L-glutamine, penicillin/streptomycin, MEM non-essential amino acids, HEPES, sodium pyruvate, β-mercaptoethanol, and murine GM-CSF (15 ng/mL; Peprotech) or human Flt3L (100 ng/mL; Peprotech). GM-CSF-derived BMDCs were used directly for all RNAi experiments unsing lentiviral shRNAs. For all other experiments, floating cells from GM-CSF cultures were sorted at day 5 by MACS using the CD11c (N418) MicroBeads kit (Miltenyi Biotec). Sorted $CD11c^+$ cells were used as GM-CSF-derived BMDCs, and plated at 10^6 cells/mL and stimulated at 16 h post sorting. For Flt3L culture, floating cells

were harvested at day 6-8 and used as Flt3L-derived BMDCs by plating them at 10^6 cells/mL and stimulating 16 h later.

For SILAC experiments, GM-CSF-derived BMDCs were grown in media containing either normal L-arginine (Arg-0) and L-lysine (Lys-0) (Sigma) or L-arginine 13C6-15N4 (Arg-10) and L-lysine 13C6-15N2 (Lys-8) (Sigma Isotec). Concentrations for L-arginine and L-lysine were 42 mg/L and 40 mg/L, respectively. The cell culture media, RPMI-1640 deficient in L-arginine and L-lysine, was a custom media preparation from Caisson Laboratories (North Logan, UT) and dialyzed serum was obtained from SAFC-Sigma. We followed all standard SILAC media preparation and labeling steps as previously described (Ong and Mann, 2006).

Preparation of primary lung fibroblasts

Mouse lung fibroblasts (MLFs) were derived from lung tissue from 6-8 week old female C57BL/6J mice (Jackson Laboratories). MLFs were isolated as previously described (Tager et al., 2004). Briefly, lungs were digested for 45 min at 37°C in collagenase and DNase I, filtered, washed, and cultured in DMEM supplemented with 15% FBS. Cells were used for experiments between passages 2 and 5.

Genetically modified mice

Bone marrow from $Plk2^{-/-}$ mice and their wild-type littermates were obtained from Elan Pharmaceuticals (Inglis et al., 2009). $Ifnar1^{-/-}$ mice on a C57BL/6J background were a gift from Kate Fitzgerald (originally from Jonathan Sprent based on Müller et al., 1994). Heterozygous $Crkl^{+/-}$ mice on a C57BL/6J background were obtained from the Jackson Laboratory. $Crkl^{+/-}$ C57BL/6J mice were crossed to wild-type Black Swiss mice from Taconic, as $Crkl^{-/-}$ mice on a pure C57BL/6J genetic background have been reported to be embryonic lethal (Guris et al., 2001; Hemmeryckx et al., 2002). Heterozygous $Crkl^{+/-}$ offspring were backcrossed to $Crkl^{+/-}$

C57BL/6J mice to obtain $Crkl^{-/-}$ mice. Mice were kept in a specific pathogen-free facility at MIT. Animal procedures were in accordance with National Institutes of Health Guidelines on animal care and use, and were approved by the MIT Committee on Animal Care (Protocol #0609-058-12).

Viruses

Sendai virus (SeV), strain Cantell, and Encephalomyocarditis virus (EMCV), strain EMC, were from ATCC. Newcastle disease virus (NDV), strain Hitchner B1 was from Aldolfo Garcia-Sastre (Mount Sinai School of Medicine), and vesicular stomatitis virus (VSV), strain Indiana was from Ulrich von Andrian (Harvard Medical School). Influenza A virus strain A/PR/8/34 and ΔNS1 were grown in Vero cells (which allow efficient growth of the ΔNS1 virus) in serum-free DMEM supplemented with 10% BSA and 1 mg/ml TPCK trypsin. Viral titers were determined by standard MDCK plaque assay. To measure the amount of VSV RNA present in infected tissues, we used previously reported qPCR primers: VSV Forward 5'-TGATACAGTACAATTATTTTGGGAC-3', and VSV Reverse 5'-GAGACTTTCTGTTACGGGATCTGG-3' (Hole et al., 2006). Viruses were handled according to CDC and NIH guidelines with protocols approved by the Broad Institutional Biosafety Committee.

Reagents

TLR ligands were from Invivogen (Pam3CSK4, ultra-pure *E. coli* K12 LPS, ODN 1585 CpG type A, and ODN 1668 CpG type B) and Enzo Life Sciences (poly(I:C)), and were used at the following concentrations: Pam3CSK4 (250 ng/mL), poly(I:C) (10 μg/mL), LPS (100 ng/mL), CpG-A (10 μg/mL), CpG-B (10 μg/mL). Heat-killed *Listeria monocytogenes* (HKLM) was from Invivogen. Polo-like kinase inhibitors were from Selleck (BI 2536; (Steegmaier et al., 2007), Sigma (GW843682X, also known as compound 1 and GSK461364; (Lansing et al., 2007), and

Chembridge (Poloxipan; (Reindl et al., 2009). SP 600125 (Jnk inhibitor) was from Enzo Life Sciences. Image-iT FX Signal Enhancer, DAPI, and Alexa Fluor Labeled Secondary Antibodies were obtained from Invitrogen. For immunofluorescence, antibodies against IRF3 (4302S) and NF-κB P65 (4764S) were obtained from Cell Signaling Technology. For cell viability assays, Alamar Blue was from Invitrogen and CellTiter-Glo from Promega.

Virus titering of MLF supernatant

293T cells were seeded and transfected with a vRNA luciferase reporter plasmid as previously described (Shapira et al., 2009). Briefly, at 24 h post-transfection, 10^4 transfected reporter cells were re-seeded in white Costar plates. Supernatants from influenza-infected MLFs were added to reporter cells and incubated for 24 h. Reporter activity was measured with firefly luciferase substrate (Steady-Glo, Promega). Luminescence activity was quantified with the Envision Multilabel Reader (Perkin Elmer).

Cell cycle analysis

Cells were fixed in ethanol, washed, and stained for 30 min at room temperature (RT) with propidium iodide (100 µg/mL) prepared in PBS (calcium- and magnesium-free) supplemented with RNAse A (2 mg/mL; Novagen) and triton X-100 (0.1%). Samples were analyzed for DNA content using an Accuri C6 flow cytometer (Accuri) and data was processed using the FlowJo software (Treestar).

ELISA

Cell culture supernatants were assayed using a sandwich ELISA kit for mouse IFN-β (PBL Biomedical Laboratories).

mRNA isolation

Total RNA was extracted with QIAzol reagent following the miRNeasy kit's procedure (Qiagen), and sample quality was tested on a 2100 Bioanalyzer (Agilent). RNA was reverse transcribed with the High Capacity cDNA Reverse Transcription kit (Applied Biosystems). For experiments with more than 12 samples, we harvested polyA+ RNA in 96- or 384-well plates with the Turbocapture mRNA kit (Qiagen) and reverse transcribed with the Sensiscript RT kit (Qiagen).

qPCR measurements

Real time quantitative PCR reactions were performed on the LightCycler 480 system (Roche) with FastStart Universal SYBR Green Master Mix (Roche). Every reaction was run in triplicate and GAPDH levels were used as an endogenous control for normalization.

shRNA knockdowns

High titer lentiviruses encoding shRNAs targeting genes of interest were obtained from The RNAi Consortium (TRC; Broad Institute, Cambridge, MA, USA) (Moffat et al., 2006). Bone marrow cells were infected with lentiviruses as described (Amit et al., 2009). For each gene of interest, we tested five shRNAs for knock down efficiency using qPCR of the target gene. We selected shRNAs with >75% knockdown efficacy. For combinatorial knockdown, two independent mixtures of two lentiviruses encoding validated shRNAs against Plk2 and 4, respectively, were used to infect bone marrow cells (two Plk2- and two Plk4-specific shRNAs were used to generate these mixtures). Lentivirus-infected cells were composed of ~90% $CD11c^+$ cells, which was comparable to sorted BMDCs and to our previous observations (Amit et al., 2009).

mRNA measurements on nCounter

Details on the nCounter system are presented in full in (Geiss et al., 2008). We used a custom CodeSet constructed to detect a total of 128 genes (including 10 control genes whose expression remain unaffected by TLR stimulation) selected by the GeneSelector algorithm (Amit et al., 2009) as described below. ~5 × 10^4 bone marrow-derived DCs were lysed in RLT buffer (Qiagen) supplemented with 1% β-mercaptoethanol. 10% of the lysate was hybridized for 16 hours with the CodeSet and loaded into the nCounter prep station followed by quantification using the nCounter Digital Analyzer following the manufacturer's instructions.

Custom Nanostring CodeSet construction using the GeneSelector algorithm

We used the CodeSet that we previously used and described in (Amit et al., 2009). Briefly, to choose a set of genes that will capture as much as possible of the information on the expression of all genes, we used an information-theoretic approach. We modeled the expression levels *X* given the experimental condition *C* with a naive Bayes model where the expression of gene *i* under condition *c* follows a normal distribution $X_i \mid C = c \sim \mathrm{N}(\mu_{ic}, \sigma_i^2)$. In this model, the expression levels of all genes depend on the experimental condition *C*, and we selected genes that are highly informative about *C*. Formally, for a set of genes *Y* we used the conditional entropy $H(C \mid \mathbf{Y}) = -\sum_c p(C=c) \sum_y p(Y=y \mid C=c) \log p(C=c \mid Y=y)$ as a measure of the remaining uncertainty in *C* once the expression levels *Y* are known. We then used this measure and a greedy procedure to select multiple disjoint gene sets, $Y_1, \ldots, Y_k$ such that for each set Y_i, $H(C \mid Y_i) < \eta$ (we set η = 0.5). In the greedy procedure, we select genes one at a time, and with each selected gene re-compute the entropy given the genes already selected in the current set. Once a set is complete (the remaining conditional entropy is below the threshold η), we add all the genes to the selected set, and repeat the procedure (excluding all the selected genes from

consideration). We stop when the number of selected genes has reached a user-defined threshold, set by the number of genes feasible for the experimental assay.

To select a time point, we used the same approach. Here, we measured entropy under all time points for multiple randomly selected gene sets of several sizes and plotted the average entropy for each timepoint (see (Amit et al., 2009)). We chose the time point with the minimal entropy (*i.e.*, 6 h post-simulation).

nCounter data analysis

After normalization by internal Nanostring controls (spike-normalization following manufacturer's instructions), we normalized the data relying on three control genes (*Ndufa7, Tbca, Tomm7*) that are the least affected by shRNAs and LPS stimulation. Next, we log-transformed the expression values (Bengtsson and Hossjer, 2006). Five signature genes (*Cxcl5, Fos, Fst, Ereg, and Egr2*) that were highly variable across control shRNA samples were removed from subsequent analysis. To score target genes whose expression is significantly affected by perturbations, we used a fold threshold corresponding to a false discovery rate (FDR) of 2%. For a given shRNA perturbation, a target gene was called as significantly affected when the ratio of the log-expression of this gene upon shRNA knockdown to the average log-expression of this gene in control shRNA samples was below (or above) a threshold (1/threshold), chosen such that, on average, no more than 2 hits (out of 128 genes in the Nanostring codeset) per control shRNA sample were called. Heatmaps and distance matrix analyses were generated using the software Gene-E (http://www.broadinstitute.org/cancer/software/GENE-E/).

Microarray hybridization and processing

For oligonucleotide microarray hybridization, 1 µg of RNA were labeled, fragmented, and hybridized to an Affymetrix Mouse Genome 430A 2.0 Array. After scanning, the expression value for each gene was calculated with RMA (Robust Multi-Array) normalization. The average

intensity difference values were normalized across the sample set. Probe sets that were absent in all samples according to Affymetrix flags were removed. All values below 50 were floored to 50.

Detection of regulated signaling genes

To identify differentially regulated signaling components (*i.e.*, kinases, phosphatases, and signaling adaptors or scaffolds), we defined regulated probesets for each condition (TLR agonist) as probesets displaying at least 1.7-fold up- or down-regulation in both duplicates of at least one time point, compared to unstimulated controls, using our previously published microarray dataset available in the NCBI Gene Expression Omnibus under the accession number GSE17721 (Amit et al., 2009). Differentially regulated probesets were intersected with lists of kinases, phosphatases, and signaling adaptors and scaffolds. These gene sets were generated combining data from publicly available databases: Panther (http://www.pantherdb.org), Gene Ontology (http://www.geneontology.org), and DAVID (http://david.abcc.ncifcrf.gov). Regulated signaling genes were hierarchically clustered using the Cluster software (Eisen et al., 1998).

Antiviral versus inflammatory gene enrichment

Genes whose expression changed upon BI 2536 treatment in microarrays were evaluated for their enrichment with genes involved in the antiviral and inflammatory programs. When multiple probesets were available for a given gene on the microarray, only the probeset with maximum value was kept for analysis. Thus, the complete microarray consisted of 14088 genes, among which 804 and 550 genes were identified as part of antiviral and inflammatory programs, respectively (Amit et al., 2009). We performed a hypergeometric test on genes whose expression changed at least 3-fold upon BI 2536 treatment compared to vehicle control

(DMSO), in either LPS or poly(I:C) samples. In addition, genes whose expression changed upon BI 2536 treatment in microarrays in response to LPS and/or poly(I:C) stimulation were analysed for enrichment of Gene Ontology (GO) processes and canonical pathways from curated databases using the Molecular Signature Databse (MSigDB; http://www.broadinstitute.org/gsea/msigdb/index.jsp).

Nanowire-mediated drug delivery and microscopy

BMDCs were plated on top of etched silicon nanowires (Si NWs) coated with small molecules (Shalek et al., 2010). After 24 hours, cells were stimulated with LPS or poly(I:C), and then fixed in 4% formaldehyde in PBS (RT, 10 min). After fixation, each sample was permeabilized with 0.25% Triton-X 100 in PBS (RT, 10 min), incubated with Image-iT FX Signal Enhancer (RT, 30 min), and then blocked with 10% goat serum and 0.25% Triton-X 100 in PBS (RT, 1 hour). After washing, the samples were placed in 3% IgG-Free BSA & 0.25% Triton-X 100 in PBS that contained primary antibodies against either IRF3 or NF-κB P65 (1:175 dilution) and then rocked overnight at 4 °C. The following day, the samples were washed with PBS and then incubated with an Alexa Fluor labeled secondary antibody (1:250 dilution) in 3% IgG-Free BSA & 0.25% Triton-X 100 in PBS (RT, 60 min). After washing with PBS, the samples were counterstained with 300 ng/mL of DAPI in PBS (RT, 30 min). For each experiment, every stimulus-molecule combination was prepared in triplicate. Once fixed, samples were imaged using an upright confocal microscope (Olympus). For each captured image, the nuclear fraction of the fluorescent protein was calculated after identifying nuclear boundaries using DAPI. Finally, distributions for this quantity across different conditions were compared using a one-way ANOVA analysis.

In vivo BI 2536 experiments in a VSV infection model

8-week old C57BL/6 male mice (from Charles River Laboratories) received 500 µg of BI 2536 (or vehicle) intravenously, and 50 µg into the footpad 3 hours before and 2 hours after infection with 10^6 pfu of VSV, as previously described (Iannacone et al., 2010), into the footpad. Mice were sacrificed 6 hours post-infection and the draining popliteal lymph nodes were harvested in RNAlater solution (Ambion) before subsequent RNA analysis. All experimental animal procedures were approved by the Institutional Animal Committees of Harvard Medical School and IDI. All infectious work was performed in designated BL2+ workspaces, in accordance with institutional guidelines, and approved by the Harvard Committee on Microbiological Safety.

MicroWestern Arrays

The MicroWestern Array (MWA) method previously described (Ciaccio et al., 2010) was modified to accommodate a larger number of lysates. The lysates were printed in a 'double-block' format with each MWA being 18 mm wide by 9 mm long. Twelve samples plus protein marker (Li-cor 928-40000) were printed with a non-contact piezoelectric arrayer (GeSiM NP2) along the top edge of the block, each block printed forty-eight times on the acrylamide gel. Electrophoresis, transfer, and antibody incubation were performed as previously described with the exception of using a modified 48-well gasket (The Gel Company MMH96) manually cut to have a larger block size in order to isolate antibodies on the nitrocellulose membrane per printed block. The antibodies used in this study were against β-ACTIN, GAPDH, β-TUBULIN, IκBα (clone L35A5), P65 (clone C22B4), STAT1, p-ABL(C-) (Y245), p-AKT (S473), p-AKT1/2/3 (T308), p-ATF2 (T71), p-ERK1/2 (T202/Y204), p-IKBALPHA (S32), p-IKKA/B (S176/180), p-IRF3 (S396), p-MAPKAPK2 (T222), p-MEK(1/2) (S217/221), p-MET (Y1234/1235), p-P38 (T180/Y182), p-P65 (S536), p-P70S6K (S371), p-P70S6K (T389), p-P90RSK (S380), p-PI3K P85(Y458) P55(Y199), p-PKCD (Y311), p-SAPK/JNK (T183/Y185), p-SEK1/MKK4 (T261), p-STAT1 (S727), p-STAT1 (Y701), p-STAT3 (S727). All antibodies were from Cell Signaling

Technology, except for β-ACTIN which was from Santa Cruz Biotechnology. Band intensities were quantified using Li-cor Odyssey analysis software (V3.0). Circles were applied to the appropriate band on the scanned image and the net intensity was calculated by subtracting the background intensity from the trimmed mean intensity of each band. The net intensity was divided by the average net intensities of β-actin to control for lysate protein concentration. Fold Change was then calculated in relation to time of inhibitor application (time zero).

Phosphotyrosine peptide analysis

Tyrosine-phosphorylated peptides were prepared using a PhosphoScan Kit (Cell Signaling Technology) as previously described (Rush et al., 2005). Briefly, 100 million cells were lysed in lysis buffer (20 mM HEPES, 25 mM sodium pyrophosphate, 10 mM beta-glycerophosphate, 9 M urea, 1 mM ortho-vanadate, 1 Roche Ser/Thr phosphatase inhibitor tablet) assisted by sonication on ice using Misonix S-4000 sonicator with five 30-second bursts at 4 watts. Lysates were pre-cleared by centrifugation for 15 min at 20,000 g. ~10 mg of total proteins from each SILAC label were mixed, reduced with 10 mM dithiothreitol and alkylated with 25 mM iodoacetamide. After 4-fold dilution 200 µg sequencing grade modified trypsin (Promega, V5113) was added in an enzyme to substrate ratio of 1:100. The total peptide mixtures were then desalted using a tC18 SepPak cartridge (Waters, 500 mg, W AT036790) and resuspended in IAP buffer (50 mM MOPS/NaOH pH 7.2, 10 mM Na2HPO4, 50 mM NaCl). Peptide immunoprecipitation was performed with protein-G agarose bead-bound anti-phosphotyrosine antibodies pY100. Peptides captured by phosphotyrosine antibodies were eluted under acidic conditions (0.15% trifluoroacetic acid). The IP eluate was analyzed by data-dependent LC-MS/MS using a Thermo LTQ-Orbitrap instrument.

Global serine, threonine, and tyrosine phosphorylation analysis

Quantitative analysis of serine, threonine and tyrosine phosphorylated peptides was performed essentially as described (Villen and Gygi, 2008) with some modifications. After stimulation, cells were lysed for 20 min in ice-cold lysis buffer (8 M Urea, 75 mM NaCl, 50 mM Tris pH 8.0, 1 mM EDTA, 2 µg/ml Aprotinin (Sigma, A6103), 10 µg/ml Leupeptin (Roche, #11017101001), 1 mM PMSF, 10 mM NaF, 2 mM Na3VO4, 50 ng/ml Calyculin A (Calbiochem, #208851), Phosphatase inhibitor cocktail 1 (1/100, Sigma, P2850) and Phosphatase inhibitor cocktail 2 (1/100, Sigma, P5726). Lysates were precleared by centrifugation at 16,500 g for 10 min and protein concentrations were determined by BCA assay (Pierce). We obtained 3 mg total protein per label out of 30 million cells. Cell lysates were mixed in equal amounts per label and proteins were reduced with 5 mM dithiothreitol and alkylated with 10 mM iodoacetamide. Samples were diluted 1:4 with HPLC water (Baker) and sequencing-grade modified trypsin (Promega, V5113) was added in an enzyme to substrate ratio of 1:150. After 16 h digest, samples were acidified with 0.5% trifluoroacetic acid (final concentration). Tryptic peptides were desalted on reverse phase tC18 SepPak columns (Waters, 500 mg, WAT036790) and lyophilized to dryness. Peptides were reconstituted in 500 µl strong cation exchange buffer A (7 mM KH2PO4, pH 2.65, 30% MeCN) and separated on a Polysulfoethyl A column from PolyLC (250 x 9.4 mm, 5 µm particle size, 200 A pore size) using an Akta Purifier 10 system (GE Healthcare). We used an 80 min gradient with a 20 min equilibration phase with buffer A, a linear increase to 30% buffer B (7 mM KH2PO4, pH 2.65, 350 mM KCL, 30% MeCN) within 33 min, 100% B for 7 min and a final equilibration with Buffer A for 20 min. The flow rate was 3 ml/min and the sample was injected after the initial 20 min equilibration phase. Upon injection, 3 ml fractions were collected with a P950 fraction collector throughout the run. 60 fractions were collected of which 3-4 adjacent fractions were combined to obtain 12 samples. Pooling of SCX fractions was guided by the UV214-trace and fractions were combined starting where the first peptide peak appeared. The 12 samples were desalted with reverse phase tC18 SepPak columns (Waters, 100 mg,

WAT036820) and lyophilized to dryness. SCX-separated peptides were subjected to IMAC (immobilized metal affinity chromatography) as described (Villen and Gygi, 2008). Briefly, peptides were reconstituted in 200 µl IMAC binding buffer (40% MeCN, 0.1% FA) and incubated for 1 h with 5 µl of packed Phos-Select beads (Sigma, P9740) in batch mode. After incubation, samples were loaded on C18 StageTips (Rappsilber et al., 2007), washed twice with 50µl IMAC binding buffer and washed once with 50µl 1% formic acid. Phosphorylated peptides were eluted from the Phos-Select resin to the C18 material by loading 3 times 70 µl of 500 mM K2HPO4 (pH 7.0). StageTips were washed with 50 µl of 1% formic acid to remove phosphate salts and eluted with 80 µl of 50% MeCN / 0.1 % formic acid. Samples were dried down by vacuum centrifugation and reconstituted in 8 µl 3% MeCN / 0.1 % formic acid.

NanoLC-MS/MS analysis

All peptide samples were separated on an online nanoflow HPLC system (Agilent 1200) and analyzed on a LTQ Orbitrap Velos (Thermo Fisher Scientific) mass spectrometer. 4 µl of peptide sample were autosampled onto a 14 cm reverse phase fused-silica capillary column (New Objective, PicoFrit PF360-75-10-N-5 with 10 µm tip opening and 75 µm inner diameter) packed in-house with 3 µm ReproSil-Pur C18-AQ media (Dr. Maisch GmbH). The HPLC setup was connected via a custom-made electrospray ion source to the mass spectrometer. After sample injection, peptides were separated at an analytical flowrate of 200 nL/min with an 70 min linear gradient (~ 0.29 %B/min) from 10% solvent A (0.1% formic acid in water) to 30% solvent B (0.1% formic acid/90% acetonitrile). The run time was 130 min for a single sample, including sample loading and column reconditioning. Data-dependent acquisition was performed using the Xcalibur 2.1 software in positive ion mode. The instrument was recalibrated in real-time by co-injection of an internal standard from ambient air (“lock mass option”) (Olsen et al., 2005). Survey spectra were acquired in the orbitrap with a resolution of 60,000 and a mass range from 350 to 1750 m/z. In parallel, up to 16 of the most intense ions per cycle were isolated,

fragmented and analyzed in the LTQ part of the instrument. Ions selected for MS/MS were dynamically excluded for 20 s after fragmentation. For the second biological replicate analysis peptides observed to be regulated in the first analysis were loaded into a global parent mass inclusion list and 4 MS/MS scans were reserved for precursors from the inclusion list whereas 12 were performed on the most intense ions per duty cycle.

Identification and quantification of peptides and proteins

Mass spectra were processed using the Spectrum Mill software package (Agilent Technologies) v4.0 b that includes in-house developed features for SILAC-based quantitation and phoshosite localization and also with the MaxQuant software package (version 1.0.13.13) (Cox and Mann, 2008), which was used in combination with a Mascot search engine (version 2.2.0, Matrix Science). For peptide identification in Spectrum Mill an International Protein Index protein sequence database (IPI version 3.60, mouse) was used which was reversed on-the-fly at search time. In MaxQuant a concatenated forward and reversed IPI protein sequence database (version 3.60, mouse) was queried. The mass tolerance for precursor ions and for fragment ions was set to 7 ppm and 0.5 Da, respectively. Cysteine carbamidomethylation was searched as a fixed modification, whereas oxidation on methionine, N-acetylation (Protein) and phosphorylation on serine, threonine or tyrosine residues were considered as variable modifications. The enzyme specificity was set to trypsin and cleavage N-terminal of proline was allowed. The maximum of missed cleavages was set to 3. For peptide identification the maximum peptide FDR was set to 1%. The minimum identification score was to 5 in Spectrum Mill and to 10 in MaxQuant. SILAC ratios were obtained from the peptide export table in Spectrum Mill and the evidence table in MaxQuant. Arginine to Proline conversion was determined to be 3.42% and 5.55% for both biological replicates, respectively. The conversion was calculated by defining Arg-10 as a fixed modification and by quantifying the ratio between peptides containing normal L-proline (Pro-0) and 13C5-15N1-labeled proline (Pro-6) with

MaxQuant. Each peptide SILAC ratio was corrected for arginine to proline conversion by the formula r[c] = r[o]/((1-p)^n), where r[c] is the corrected ratio, r[o] the observed ratio, p the conversion rate and n the number of proline residues per peptide. The median ratios of all non-phosphorylated peptides were used to normalize the M/L and H/L ratios of all phosphorylated peptides. To allow better peptide grouping, phosphosite localization information obtained from SpectrumMill and MaxQuant were further simplified. Probability scores greater or equal 0.75 were called fully localized and designated with (1.0), scores smaller 0.75 and greater or equal to 0.5 were called ambiguously localized and designated with (0.5), whereas scores smaller than 0.5 were called non-localized and the total number of phosphorylation sites per peptide was designated with an underscore after the peptide sequence. Median SILAC ratios of phosphopeptides for each experiment were calculated over all versions of the same peptide including different charge states and methionine oxidation states. The highest scoring versions of each distinct peptide were reported per experiment. Overlapping data between SpectrumMill and MaxQuant as well as between different biological replicates was analyzed for discrepancies by calculating the mean and standard deviation over all residuals for different ratios of the same phosphopeptide. Residuals were calculated by subtracting the two values for each phosphopeptide derived by SpectrumMill or MaxQuant as well as by two different biological replicates. All peptides were filtered from the data set that had residuals greater than 3 standard deviations distant from the mean as they were not reproducible between two biological replicates or between SpectrumMill and MaxQuant. Data derived from both software packages was combined and MaxQuant data was reported when the same phosphopeptide was identified and quantified by both programs. Log2 phosphopeptide ratios of BI-2536 treated vs untreated dendritic cells followed a normal distribution that was fitted using least squares regression. Mean and standard deviation values derived from the Gaussian fit were used to calculate p-values. An FDR-based measure was used to assess significance of the findings (Storey and Tibshirani, 2003).

Accession numbers and supplementary tables

Complete microarray data sets are available in the NCBI Gene Expression Omnibus under the accession number GSE28520. The proteomics raw data (.raw files) associated with this manuscript were submitted to the Tranche data repository and can be downloaded from: https://proteomecommons.org/tranche/, using the following hash:

HTWY5ZeSLM1hyYEyfEiJREkgLXs6BZxCczuixy9XjULsync5HCkXx/8gB7nZKpGocwOnt8vOk/

Q3cpbPh/ycD/2LT0AAAAAAAAAuEg==

and passphrase:

SpSTB6vceSUKeNqefq59

All 7 supplementary tables containing the datasets presented in this Chapter are available with the online version of this published work (Chevrier et al., 2011).

3.14. References

Amit, I., Citri, A., Shay, T., Lu, Y., Katz, M., Zhang, F., Tarcic, G., Siwak, D., Lahad, J., Jacob-Hirsch, J., et al. (2007). A module of negative feedback regulators defines growth factor signaling. Nat Genet *39*, 503-512.

Amit, I., Garber, M., Chevrier, N., Leite, A.P., Donner, Y., Eisenhaure, T., Guttman, M., Grenier, J.K., Li, W., Zuk, O., et al. (2009). Unbiased reconstruction of a mammalian transcriptional network mediating pathogen responses. Science *326*, 257-263.

Archambault, V., and Glover, D.M. (2009). Polo-like kinases: conservation and divergence in their functions and regulation. Nat Rev Mol Cell Biol *10*, 265-275.

Ballif, B.A., Carey, G.R., Sunyaev, S.R., and Gygi, S.P. (2008). Large-scale identification and evolution indexing of tyrosine phosphorylation sites from murine brain. J Proteome Res *7*, 311-318.

Banchereau, J., and Pascual, V. (2006). Type I interferon in systemic lupus erythematosus and other autoimmune diseases. Immunity *25*, 383-392.

Barbie, D.A., Tamayo, P., Boehm, J.S., Kim, S.Y., Moody, S.E., Dunn, I.F., Schinzel, A.C., Sandy, P., Meylan, E., Scholl, C., et al. (2009). Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature *462*, 108-112.

Barrat, F.J., and Coffman, R.L. (2008). Development of TLR inhibitors for the treatment of autoimmune diseases. Immunol Rev *223*, 271-283.

Bengtsson, H., and Hossjer, O. (2006). Methodological study of affine transformations of gene expression data with proposed robust non-parametric multi-dimensional normalization method. BMC Bioinformatics *7*, 100.

Birge, R.B., Kalodimos, C., Inagaki, F., and Tanaka, S. (2009). Crk and CrkL adaptor proteins: networks for physiological and pathological signaling. Cell Commun Signal *7*, 13.

Blasius, A.L., and Beutler, B. (2010). Intracellular toll-like receptors. Immunity *32*, 305-315.

Chang, J., Cizmecioglu, O., Hoffmann, I., and Rhee, K. (2010). PLK2 phosphorylation is critical for CPAP function in procentriole formation during the centrosome cycle. EMBO J *29*, 2395-2406.

Chien, Y., Kim, S., Bumeister, R., Loo, Y.M., Kwon, S.W., Johnson, C.L., Balakireva, M.G., Romeo, Y., Kopelovich, L., Gale, M., Jr., et al. (2006). RalB GTPase-mediated activation of the IkappaB family kinase TBK1 couples innate immune signaling to tumor cell survival. Cell *127*, 157-170.

Chevrier N, Mertins P, Artyomov MN, Shalek AK, Iannacone M, Ciaccio MF, Gat-Viks I, Tonti E, DeGrace MM, Clauser KR, et al. (2011). Systematic Discovery of Signaling Components Identifies New Branches in Viral-Sensing Pathways. Cell *147* 853-67.

Chu, W.M., Ostertag, D., Li, Z.W., Chang, L., Chen, Y., Hu, Y., Williams, B., Perrault, J., and Karin, M. (1999). JNK2 and IKKbeta are required for activating the innate response to viral infection. Immunity *11*, 721-731.

Ciaccio, M.F., Wagner, J.P., Chuu, C.P., Lauffenburger, D.A., and Jones, R.B. (2010). Systems analysis of EGF receptor signaling dynamics with microwestern arrays. Nat Methods *7*, 148-155.

Cizmecioglu, O., Warnke, S., Arnold, M., Duensing, S., and Hoffmann, I. (2008). Plk2 regulated centriole duplication is dependent on its localization to the centrioles and a functional polo-box domain. Cell Cycle *7*, 3548-3555.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol *26*, 1367-1372.

Crozat, K., Hoebe, K., Ugolini, S., Hong, N.A., Janssen, E., Rutschmann, S., Mudd, S., Sovath, S., Vivier, E., and Beutler, B. (2007). Jinx, an MCMV susceptibility phenotype caused by disruption of Unc13d: a mouse model of type 3 familial hemophagocytic lymphohistiocytosis. J Exp Med *204*, 853-863.

Doyle, S., Vaidya, S., O'Connell, R., Dadgostar, H., Dempsey, P., Wu, T., Rao, G., Sun, R., Haberland, M., Modlin, R., et al. (2002). IRF3 mediates a TLR3/TLR4-specific antiviral gene program. Immunity *17*, 251-263.

Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A *95*, 14863-14868.

Fraser, I.D., and Germain, R.N. (2009). Navigating the network: signaling cross-talk in hematopoietic cells. Nat Immunol *10*, 327-331.

Freeman, M. (2000). Feedback control of intercellular signalling in development. Nature *408*, 313-319.

Geiss, G.K., Bumgarner, R.E., Birditt, B., Dahl, T., Dowidar, N., Dunaway, D.L., Fell, H.P., Ferree, S., George, R.D., Grogan, T., et al. (2008). Direct multiplexed measurement of gene expression with color-coded probe pairs. Nat Biotechnol *26*, 317-325.

Guris, D.L., Fantes, J., Tara, D., Druker, B.J., and Imamoto, A. (2001). Mice lacking the homologue of the human 22q11.2 gene CRKL phenocopy neurocristopathies of DiGeorge syndrome. Nat Genet *27*, 293-298.

Heikkinen, L.S., Kazlauskas, A., Melen, K., Wagner, R., Ziegler, T., Julkunen, I., and Saksela, K. (2008). Avian and 1918 Spanish influenza a virus NS1 proteins bind to Crk/CrkL Src homology 3 domains to activate host cell signaling. J Biol Chem *283*, 5719-5727.

Hemmeryckx, B., Reichert, A., Watanabe, M., Kaartinen, V., de Jong, R., Pattengale, P.K., Groffen, J., and Heisterkamp, N. (2002). BCR/ABL P190 transgenic mice develop leukemia in the absence of Crkl. Oncogene *21*, 3225-3231.

Hennessy, E.J., Parker, A.E., and O'Neill, L.A. (2010). Targeting Toll-like receptors: emerging therapeutics? Nat Rev Drug Discov *9*, 293-307.

Hole, K., Clavijo, A., and Pineda, L.A. (2006). Detection and serotype-specific differentiation of vesicular stomatitis virus using a multiplex, real-time, reverse transcription-polymerase chain reaction assay. J Vet Diagn Invest *18*, 139-146.

Hrincius, E.R., Wixler, V., Wolff, T., Wagner, R., Ludwig, S., and Ehrhardt, C. (2010). CRK adaptor protein expression is required for efficient replication of avian influenza A viruses and controls JNK-mediated apoptotic responses. Cell Microbiol *12*, 831-843.

Iannacone, M., Moseman, E.A., Tonti, E., Bosurgi, L., Junt, T., Henrickson, S.E., Whelan, S.P., Guidotti, L.G., and von Andrian, U.H. (2010). Subcapsular sinus macrophages prevent CNS invasion on peripheral infection with a neurotropic virus. Nature *465*, 1079-1083.

Inglis, K.J., Chereau, D., Brigham, E.F., Chiou, S.S., Schobel, S., Frigon, N.L., Yu, M., Caccavello, R.J., Nelson, S., Motter, R., et al. (2009). Polo-like kinase 2 (PLK2) phosphorylates alpha-synuclein at serine 129 in central nervous system. J Biol Chem *284*, 2598-2602.

Kanzler, H., Barrat, F.J., Hessel, E.M., and Coffman, R.L. (2007). Therapeutic targeting of innate immunity with Toll-like receptor agonists and antagonists. Nat Med *13*, 552-559.

Kato, H., Sato, S., Yoneyama, M., Yamamoto, M., Uematsu, S., Matsui, K., Tsujimura, T., Takeda, K., Fujita, T., Takeuchi, O., et al. (2005). Cell type-specific involvement of RIG-I in antiviral response. Immunity *23*, 19-28.

Kato, H., Takeuchi, O., Sato, S., Yoneyama, M., Yamamoto, M., Matsui, K., Uematsu, S., Jung, A., Kawai, T., Ishii, K.J., et al. (2006). Differential roles of MDA5 and RIG-I helicases in the recognition of RNA viruses. Nature *441*, 101-105.

Kawagoe, T., Takeuchi, O., Takabatake, Y., Kato, H., Isaka, Y., Tsujimura, T., and Akira, S. (2009). TANK is a negative regulator of Toll-like receptor signaling and is critical for the prevention of autoimmune nephritis. Nat Immunol *10*, 965-972.

Keller, A.D., and Maniatis, T. (1991). Identification and characterization of a novel repressor of beta-interferon gene expression. Genes Dev *5*, 868-879.

Kuan, A.P., Chamberlain, W., Malkiel, S., Lieu, H.D., Factor, S.M., Diamond, B., and Kotzin, B.L. (1999). Genetic control of autoimmune myocarditis mediated by myosin-specific antibodies. Immunogenetics *49*, 79-85.

Laloraya, M., Davoodi-Semiromi, A., Kumar, G.P., McDuffie, M., and She, J.X. (2006). Impaired Crkl expression contributes to the defective DNA binding of Stat5b in nonobese diabetic mice. Diabetes *55*, 734-741.

Lande, R., Gregorio, J., Facchinetti, V., Chatterjee, B., Wang, Y.H., Homey, B., Cao, W., Su, B., Nestle, F.O., Zal, T., et al. (2007). Plasmacytoid dendritic cells sense self-DNA coupled with antimicrobial peptide. Nature *449*, 564-569.

Lansing, T.J., McConnell, R.T., Duckett, D.R., Spehar, G.M., Knick, V.B., Hassler, D.F., Noro, N., Furuta, M., Emmitte, K.A., Gilmer, T.M., et al. (2007). In vitro biological activity of a novel small-molecule inhibitor of polo-like kinase 1. Mol Cancer Ther *6*, 450-459.

Luber, C.A., Cox, J., Lauterbach, H., Fancke, B., Selbach, M., Tschopp, J., Akira, S., Wiegand, M., Hochrein, H., O'Keeffe, M., et al. (2010). Quantitative proteomics reveals subset-specific viral recognition in dendritic cells. Immunity *32*, 279-289.

Mandl, J.N., Barry, A.P., Vanderford, T.H., Kozyr, N., Chavan, R., Klucking, S., Barrat, F.J., Coffman, R.L., Staprans, S.I., and Feinberg, M.B. (2008). Divergent TLR7 and TLR9 signaling and type I interferon production distinguish pathogenic and nonpathogenic AIDS virus infections. Nat Med *14*, 1077-1087.

Matsui, K., Kumagai, Y., Kato, H., Sato, S., Kawagoe, T., Uematsu, S., Takeuchi, O., and Akira, S. (2006). Cutting edge: Role of TANK-binding kinase 1 and inducible IkappaB kinase in IFN responses against viruses in innate immune cells. J Immunol *177*, 5785-5789.

Moffat, J., Grueneberg, D.A., Yang, X., Kim, S.Y., Kloepfer, A.M., Hinkle, G., Piqani, B., Eisenhaure, T.M., Luo, B., Grenier, J.K., et al. (2006). A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. Cell *124*, 1283-1298.

Mross, K., Frost, A., Steinbild, S., Hedbom, S., Rentschler, J., Kaiser, R., Rouyrre, N., Trommeshauser, D., Hoesl, C.E., and Munzert, G. (2008). Phase I dose escalation and pharmacokinetic study of BI 2536, a novel Polo-like kinase 1 inhibitor, in patients with advanced solid tumors. J Clin Oncol *26*, 5511-5517.

Muller, U., Steinhoff, U., Reis, L.F., Hemmi, S., Pavlovic, J., Zinkernagel, R.M., and Aguet, M. (1994). Functional role of type I and type II interferons in antiviral defense. Science *264*, 1918-1921.

Oganesyan, G., Saha, S.K., Guo, B., He, J.Q., Shahangian, A., Zarnegar, B., Perry, A., and Cheng, G. (2006). Critical role of TRAF3 in the Toll-like receptor-dependent and -independent antiviral response. Nature *439*, 208-211.

Olsen, J.V., de Godoy, L.M., Li, G., Macek, B., Mortensen, P., Pesch, R., Makarov, A., Lange, O., Horning, S., and Mann, M. (2005). Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. Mol Cell Proteomics *4*, 2010-2021.

Ong, S.E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A., and Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics *1*, 376-386.

Ong, S.E., and Mann, M. (2006). A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). Nat Protoc *1*, 2650-2660.

Rappsilber, J., Mann, M., and Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat Protoc *2*, 1896-1906.

Reindl, W., Yuan, J., Kramer, A., Strebhardt, K., and Berg, T. (2009). A pan-specific inhibitor of the polo-box domains of polo-like kinases arrests cancer cells in mitosis. Chembiochem *10*, 1145-1148.

Reizis, B., Bunin, A., Ghosh, H.S., Lewis, K.L., and Sisirak, V. (2011). Plasmacytoid dendritic cells: recent progress and open questions. Annu Rev Immunol *29*, 163-183.

Rothlin, C.V., Ghosh, S., Zuniga, E.I., Oldstone, M.B., and Lemke, G. (2007). TAM receptors are pleiotropic inhibitors of the innate immune response. Cell *131*, 1124-1136.

Rush, J., Moritz, A., Lee, K.A., Guo, A., Goss, V.L., Spek, E.J., Zhang, H., Zha, X.M., Polakiewicz, R.D., and Comb, M.J. (2005). Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. Nat Biotechnol *23*, 94-101.

Santiago-Raber, M.L., Baccala, R., Haraldsson, K.M., Choubey, D., Stewart, T.A., Kono, D.H., and Theofilopoulos, A.N. (2003). Type-I interferon receptor deficiency reduces lupus-like disease in NZB mice. J Exp Med *197*, 777-788.

Seeburg, D.P., Pak, D., and Sheng, M. (2005). Polo-like kinases in the nervous system. Oncogene *24*, 292-298.

Shalek, A.K., Robinson, J.T., Karp, E.S., Lee, J.S., Ahn, D.R., Yoon, M.H., Sutton, A., Jorgolli, M., Gertner, R.S., Gujral, T.S., et al. (2010). Vertical silicon nanowires as a universal platform for delivering biomolecules into living cells. Proc Natl Acad Sci U S A *107*, 1870-1875.

Shapira, S.D., Gat-Viks, I., Shum, B.O., Dricot, A., de Grace, M.M., Wu, L., Gupta, P.B., Hao, T., Silver, S.J., Root, D.E., et al. (2009). A physical and regulatory map of host-influenza interactions reveals pathways in H1N1 infection. Cell *139*, 1255-1267.

Sharma, K., Kumar, C., Keri, G., Breitkopf, S.B., Oppermann, F.S., and Daub, H. (2010). Quantitative analysis of kinase-proximal signaling in lipopolysaccharide-induced innate immune response. J Proteome Res *9*, 2539-2549.

Steegmaier, M., Hoffmann, M., Baum, A., Lenart, P., Petronczki, M., Krssak, M., Gurtler, U., Garin-Chesa, P., Lieb, S., Quant, J., et al. (2007). BI 2536, a potent and selective inhibitor of polo-like kinase 1, inhibits tumor growth in vivo. Curr Biol *17*, 316-322.

Storey, J.D., and Tibshirani, R. (2003). Statistical significance for genomewide studies. Proc Natl Acad Sci U S A *100*, 9440-9445.

Strebhardt, K. (2010). Multifaceted polo-like kinases: drug targets and antitargets for cancer therapy. Nat Rev Drug Discov *9*, 643-660.

Tager, A.M., Kradin, R.L., LaCamera, P., Bercury, S.D., Campanella, G.S., Leary, C.P., Polosukhin, V., Zhao, L.H., Sakamoto, H., Blackwell, T.S., et al. (2004). Inhibition of pulmonary fibrosis by the chemokine IP-10/CXCL10. Am J Respir Cell Mol Biol *31*, 395-404.

Takayanagi, H., Kim, S., Matsuo, K., Suzuki, H., Suzuki, T., Sato, K., Yokochi, T., Oda, H., Nakamura, K., Ida, N., et al. (2002). RANKL maintains bone homeostasis through c-Fos-dependent induction of interferon-beta. Nature *416*, 744-749.

Takeuchi, O., and Akira, S. (2010). Pattern recognition receptors and inflammation. Cell *140*, 805-820.

Ulevitch, R.J. (2004). Therapeutics targeting the innate immune system. Nat Rev Immunol *4*, 512-520.

Villen, J., and Gygi, S.P. (2008). The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry. Nat Protoc *3*, 1630-1638.

Vitour, D., Dabo, S., Ahmadi Pour, M., Vilasco, M., Vidalain, P.O., Jacob, Y., Mezel-Lemoine, M., Paz, S., Arguello, M., Lin, R., et al. (2009). Polo-like kinase 1 (PLK1) regulates interferon (IFN) induction by MAVS. J Biol Chem *284*, 21797-21809.

Wang, L., Gordon, R.A., Huynh, L., Su, X., Park Min, K.H., Han, J., Arthur, J.S., Kalliolias, G.D., and Ivashkiv, L.B. (2010). Indirect inhibition of Toll-like receptor and type I interferon responses by ITAM-coupled receptors and integrins. Immunity *32*, 518-530.

Weintz, G., Olsen, J.V., Fruhauf, K., Niedzielska, M., Amit, I., Jantsch, J., Mages, J., Frech, C., Dolken, L., Mann, M., et al. (2010). The phosphoproteome of toll-like receptor-activated macrophages. Mol Syst Biol *6*, 371.

Wellcome Trust Case Control Consortium (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature *447*, 661-678.

Xavier, R.J., and Rioux, J.D. (2008). Genome-wide association studies: a new window into immune-mediated diseases. Nat Rev Immunol *8*, 631-643.

Xiao, N., Eidenschenk, C., Krebs, P., Brandl, K., Blasius, A.L., Xia, Y., Khovananth, K., Smart, N.G., and Beutler, B. (2009). The Tpl2 mutation Sluggish impairs type I IFN production and increases susceptibility to group B streptococcal disease. J Immunol *183*, 7975-7983.

Ye, J., Chen, S., and Maniatis, T. (2011). Cardiac glycosides are potent inhibitors of interferon-beta gene expression. Nat Chem Biol *7*, 25-33.

Zhang, W., Wang, J., Zhang, Y., Yuan, Y., Guan, W., Jin, C., Chen, H., Wang, X., Yang, X., and He, F. (2010). The scaffold protein TANK/I-TRAF inhibits NF-kappaB activation by recruiting polo-like kinase 1. Mol Biol Cell *21*, 2500-2513.

Chapter 4 – Integrative Analysis of Signaling-to-Transcription Events Reveals the Organization of the TLR Network

4.1. Author Contributions

- Nicolas Chevrier conceived the study, designed and performed all experiments, analyzed and interpreted data, and wrote the manuscript draft.
- Philipp Mertins optimized the phosphoproteomics protocol, assisted with designing and optimizing the affinity purification protocol, assisted with execution of proteomics experiments, operated the mass spectrometer, and processed raw data files.
- Nir Yosef contributed to the statistical analyses of the following data sets: proteomics, NanoString, and ChIP-seq (from an independent study).
- Karl Clauser assisted with computational analysis of the mass spectrometry raw data sets.
- Ido Amit assisted with the design of the temporal phosphoproteomics profiling.
- Aviv Regev supervised computational analyses.
- Steven A. Carr supervised the mass spectrometry-based work.
- Nir Hacohen supervised the project.

4.2. Introduction

The studies reported in Chapter 3 demonstrate how a systematic strategy – combining transcriptomics, genetic and chemical perturbations, and unbiased phosphoproteomics – can be used to discover and position new components of signaling networks. In particular, we relied on changes in mRNA levels upon stimulation to identify candidate regulators. Although this

approach can lead to significant advances in deciphering signaling networks, it is far from exhaustive. Indeed, components with unchanged mRNA levels are missed; direct enzyme-substrate relationships cannot be established; interactions from components can be inferred from co-phenotypes (*i.e.*, similar knockdown effect on the TLR gene signature), but remain to be tested; and (4) the mechanistic links within and between the signaling and transcriptional regulatory layers are not established. To address start addressing these limitations, we reasoned that combination of our perturbation-based approach (**Chapter 2 and 3**) with large-scale biochemical measurements (*e.g.*, post-translational modifications, protein-protein interactions), will lead to more comprehensive, integrated maps of signaling and transcriptional networks (**Figure 4.1**). The two main goals of this study are: (*i*) to uncover modules of signaling regulators controlling TLR-dependent gene expression, and (*ii*) to extract organizational rules driving the 'signaling-to-transcriptional' TLR network by intersecting functional and physical interactions among the signaling and transcriptional regulatory layers.

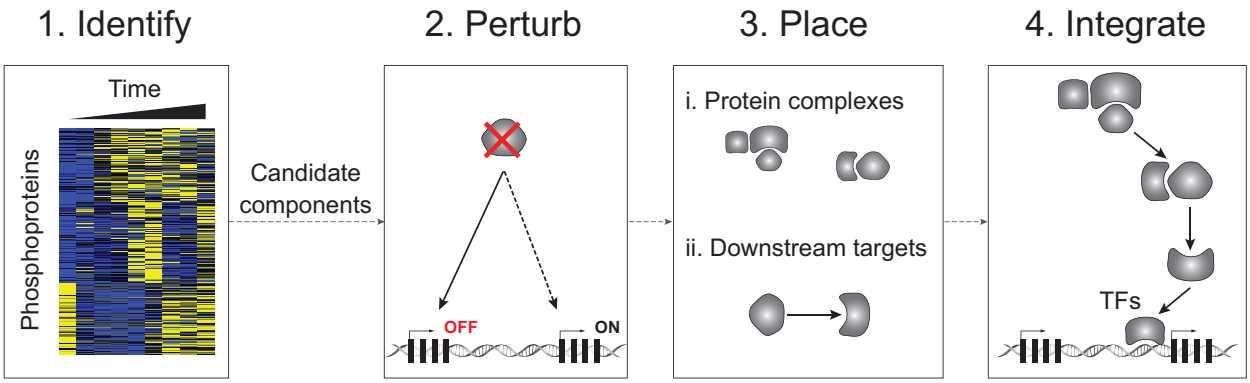

Figure 4.1. An Integrated Approach to Dissect Intracellular Networks from Signaling to Transcriptional Events

Shown is a schematic depicting the strategy consisting of four steps (from left to right): (1) extract candidate signaling regulators from temporal phosphoproteomic profiles; (2) perturb candidates with shRNAs and measure the effect on the expression of signature genes; (3) place each regulator within their functional context by dissecting protein complexes and kinase-substrate relationships; (4) integrate orthogonal datasets into a dynamic, physical model connecting signaling (*e.g.*, phosphorylation, binding) and downstream transcriptional events (*e.g.*, TF binding to genomic loci).

4.3. A Temporal Phosphoproteomic Analysis Reveals Numerous Known and Candidate Regulators of the TLR Pathways

We reasoned that large-scale measurements of the changes in protein phosphorylation in DCs would help characterize the TLR signaling network (Choudhary and Mann, 2010). Our rationale was two-fold: (1) TLR signaling has been shown to rely on multiple kinases (*e.g.*, MAP kinases, IRAKs, IKKs, TBK1, IKK-ε) and phosphorylation of key transcription factors (*e.g.*, NF-κB and IRF families, Jun, Fos) (Takeuchi and Akira, 2010), and (2) we (Chevrier et al., 2011) and others (Weintz et al., 2010) have previously shown that phosphoproteomic measurements in DCs or macrophages treated with TLR agonist allowed the identification of functional components of the TLR system. We measured large-scale changes in phosphorylation in a time course of LPS-stimulated DCs using SILAC for quantification, leading to the identification of 61299 phosphopeptides corresponding to 29783 phosphosites on 9648 distinct proteins (False Discovery Rate (FDR) 1%) (**Figure 4.2A and Experimental Procedures**). For all identical phosphosites occurring in different peptides (*e.g.*, trypsin miscleavage), we collapsed the values by averaging the data points measured more than once. To identify differentially expressed phosphopeptides at each time point, we computed the fold-change values compared to unstimulated samples and used for subsequent analyses the extreme points of the resulting Gaussian distributions ($p < 0.001$, **Figure 4.2B**). We identified 32150 and 12332 phosphopeptides with a significant fold-change in at least one time point and two consecutive time points, respectively. These phosphopeptides matched 7928 and 3194 proteins, respectively.

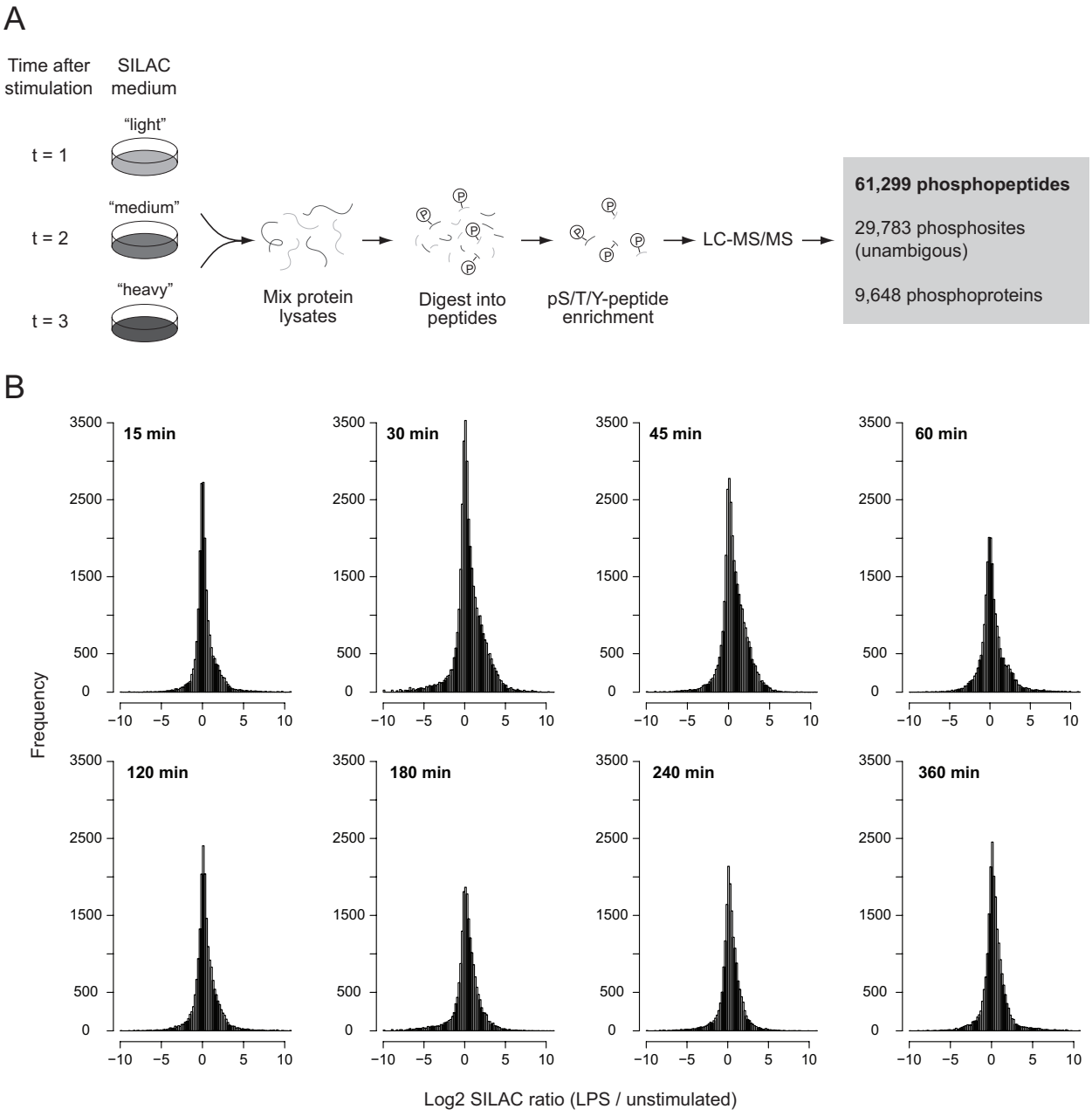

Figure 4.2. Identification of the LPS-Regulated Phosphoproteome in DCs

(A) Overview of the experimental strategy to measure large-scale changes in phosphorylation. GM-CSF-induced BMDCs were grown in different SILAC media (triplex experiments), and lysed upon stimulation with LPS at different time points. Protein lysates were mixed (1:1:1 ratio), digested into peptides with trypsin, and processed for phosphopeptide enrichment using SCX/IMAC (see Experimental Procedures) before LC-MS/MS analysis. Numbers on the right are (grey box), from top to bottom: the total number of phosphopeptides identified, the total number of phosphosites, and the total number of corresponding phosphoproteins.

(B) Frequency plots showing the distribution of the SILAC ratios (LPS/unstimulated) for each of the 8 time points measured after LPS stimulation.

Clustering analysis of the temporal data sets revealed interesting features that we exploited to identify the pathways and biological processes and functions that are enriched, and to select candidate regulators for functional characterization. To avoid potential artifacts, we focused our clustering analysis on the 3627 phosphopeptides (matching 1277 proteins) that satisfied the following criteria: (1) significantly regulated in two consecutive time points; (2) quantified in at least 6 time points out of the 8 measured; and (3) matching a protein whose mRNA was expressed based on microarray measurements. We used *k*-means clustering to partition the temporal profiles of these 3627 phosphopeptides into 15 clusters, which was the minimal *k* number of clusters providing optimal similarity (Euclidean metric) within each cluster (**Figure 4.3A**).

Enrichment analyses of the pathways and molecular functions significantly over-represented ($p < 10^{-5}$) among these 1277 LPS-regulated proteins revealed canonical members of pathogen-sensing pathways (TLR, $p = 5.96 \times 10^{-7}$; NOD-like receptor, $p = 2.75 \times 10^{-7}$; RIG-I-like receptor, $p = 2.19 \times 10^{-7}$), and associated pathways such as MAP kinases ($p = 5.78 \times 10^{-9}$) (**Figure 4.3B**). In addition, molecular functions associated with signal transduction were highly enriched, including protein kinase activity ($p = 6.33 \times 10^{-13}$), transcription factor binding ($p = 9.5 \times 10^{-8}$), and GTPases and their regulators (*e.g.*, GTPase regulator activity, $p = 2.3 \times 10^{-25}$; small GTPase regulator activity, $p = 1.23 \times 10^{-14}$) (**Figure 4.3B**). Other enriched gene sets pointed to nascent areas of TLR biology (Barton and Kagan, 2009; Lee et al., 2012), such as the organization and regulation of the TLR system within the framework of intracellular organelles and structures (*e.g.*, endocytosis, $p = 6.59 \times 10^{-6}$; cytoskeletal protein binding, $p = 1.03 \times 10^{-10}$; phospholipid binding, $p = 3.79 \times 10^{-6}$). Interestingly, comparing changes in phosphorylation (**Figure 4.3C**) and mRNA levels (**Figure 4.3D**) for known TLR pathway members suggested a more comprehensive coverage of the TLR pathways using phosphoproteomics. Altogether,

these results indicate the predictive potential of the data in identifying functional components of the TLR network.

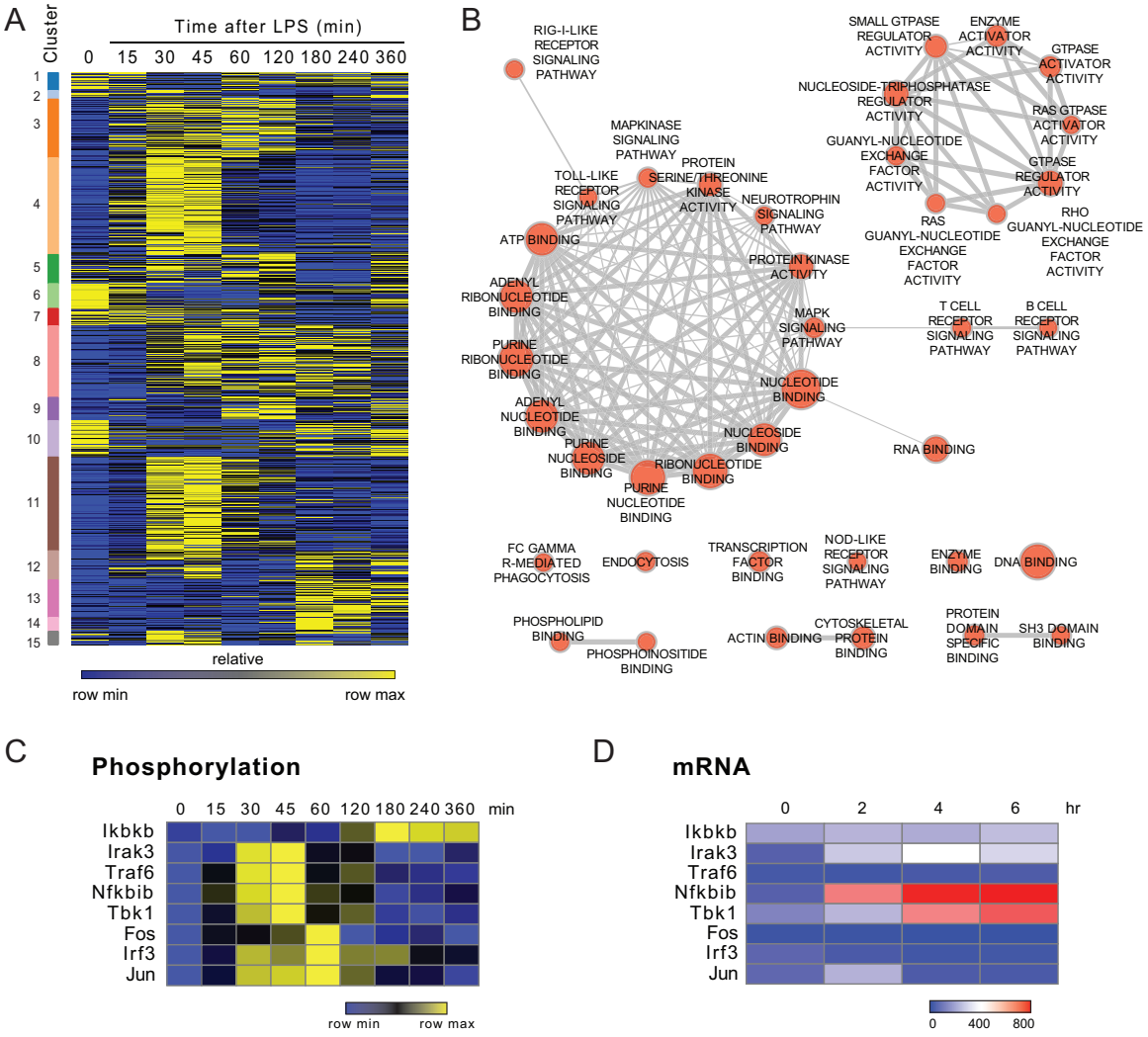

Figure 4.3. Clustering and Enrichment Analyses Reveals the Predictive Potential of the TLR Phosphoproteome in Identifying Regulators

(A) Heatmap of all the phosphopeptides (row) significantly regulated (up or down) in at least 2 consecutive time points (column). Data were clustered using *k*-means (left-most color-coded clusters), and only phosphopeptides measured in at least 6 out of 8 time points are shown. Missing values were imputed as the average between the two surrounding values), and ratios at time point 0 were set to 1 for clustering.

(B) Gene-set enrichment analysis on the 1277 phosphoproteins belonging to the 15 *k*-means clusters from (A). Significantly enriched genesets (GO terms, KEGG pathways; *p* value < 10^{-5}) are represented in a network-based view whereby each set is a node and edges represent gene overlap between sets. The size of nodes and edges is proportional to geneset sizes and the degrees of overlap, respectively. Enrichments were generated using DAVID, and the network view was created in Cytoscape using the 'Enrichment Map' plugin (Experimental Procedures).

(C-D) Examples of temporal phosphorylation (C) and mRNA (D) profiles of selected known TLR pathway members. Rows correspond to phosphopeptides (C) and mRNA transcripts (D) from indicated genes, and columns to time point after LPS stimulation.

4.4. A Perturbation Strategy Uncovers Components Participating in the Control of TLR Signature Gene Expression

The results of these temporal (**Figure 4.3A**) and enrichment (**Figure 4.3B**) analyses were used to focus the scope of the subsequent identification and characterizations of candidate signaling regulators. First, we focused our attention on the four largest *k* clusters in numbers of phosphopeptides: *k* clusters 3, 4, 8, and 11, collectively containing ~36% of all phosphoproteins present in our dataset (**Figure 4.4A**). These four clusters encompass phosphopeptides with an early peak in induction at 30 min after LPS stimulation, which is not due to changes in total protein levels that remained largely unchanged during the course of stimulation (**Figure 4.4A**). We reasoned that focusing on these early clusters would favor the identification of candidate regulators directly downstream of TLR4, thus avoiding feedbacks through transcription and autocrine/paracrine signaling. Furthermore, practically speaking, focusing on a single, early time point (*i.e.*, 30 min), helps narrowing down the scope of our subsequent functional assays (detailed below). Using molecular function annotations, we selected from these four *k* clusters (2033 phosphopeptides corresponding to 833 proteins) all 203 enzymes (*e.g.*, kinases, GTPases) and enzyme binders and regulators (*e.g.*, GTPase regulators) as candidate regulators of the TLR system (**Figure 4.4B**).

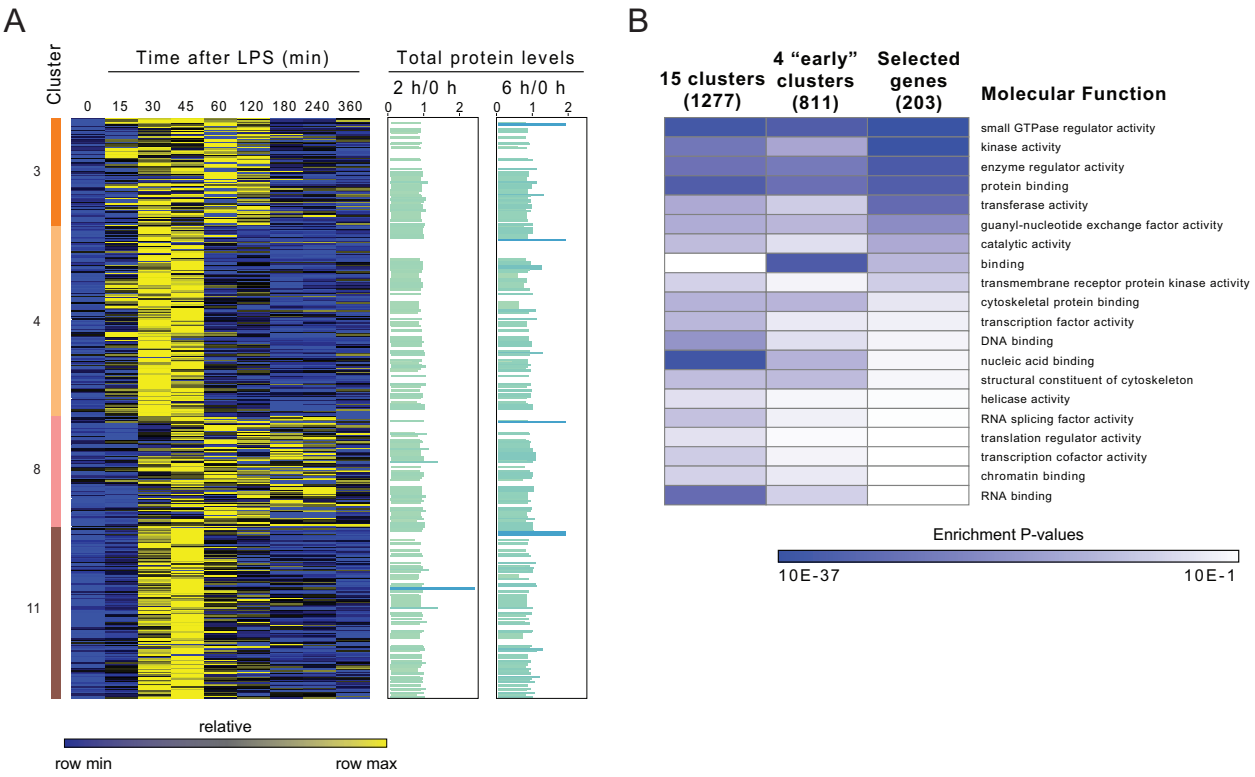

Figure 4.4. Candidate Signaling Regulators Were Selected Based on Molecular Function Annotations

(A) Heatmap of the phosphopeptides belonging to clusters 3, 4, 8, and 11 (left-most bar) as represented in **Figure 4.3A**, and matching the selected 203 candidate regulators. Bar graphs on the right indicate the total protein levels upon LPS stimulation for 2 and 6 h (as SILAC ratios between 2/0 h, and 6/0 h, from left to right; a ratio of 1 indicates no change in protein level).
(B) Functional enrichment of the molecular functions present in the proteins from all 15 *k*-means clusters (comprising 1277 phosphoproteins; left column), from *k*-means clusters 3, 4, 8, and 11 (comprising 811 phosphoproteins; middle column), and from selected candidate regulators (comprising 203 phosphoproteins; right column). Numbers in parenthesis on the top of each column indicate the total number of proteins in each case. The color-coded (blue) legend indicates *p*-values (hypergeometric test).

To assess the function of these 203 phosphoproteins, we measured the impact of genetic perturbations on a representative set of 230 TLR-induced genes, referred to as TLR signature genes (**Figure 4.5**). Out of the 203 candidates initially selected, 156 were efficiently knocked down using shRNAs, among which 30 led to a significant decrease in TLR signature genes (**Figure 4.5**). These 30 phosphoproteins included well-described TLR signaling components, such as TRIF, TBK1, and IRAK2, as well as a host of phosphoproteins not previously associated with the TLR system. Interestingly, known and candidate components had similar effects on the TLR gene signature (**Figure 4.6A**). By measuring the pairwise similarity (Pearson

correlation) among these 30 perturbation profiles, we observed three clusters of signaling regulators (**Figure 4.6B**). These genetic interactions suggest that the TLR4 pathways leading to transcriptional regulation are likely to be organized around at least three overlapping pathways or modules. Indeed, while some genes such as TRIF (*Ticam1*) led a near-complete abrogation of TLR signature gene induction, MYD88 only affected inflammatory gene expression (*e.g.*, *Cxcl1*, *Cxcl2*, *Il1a*, *Il1b*, *Il6*, *Tnf*, *Il12b*).

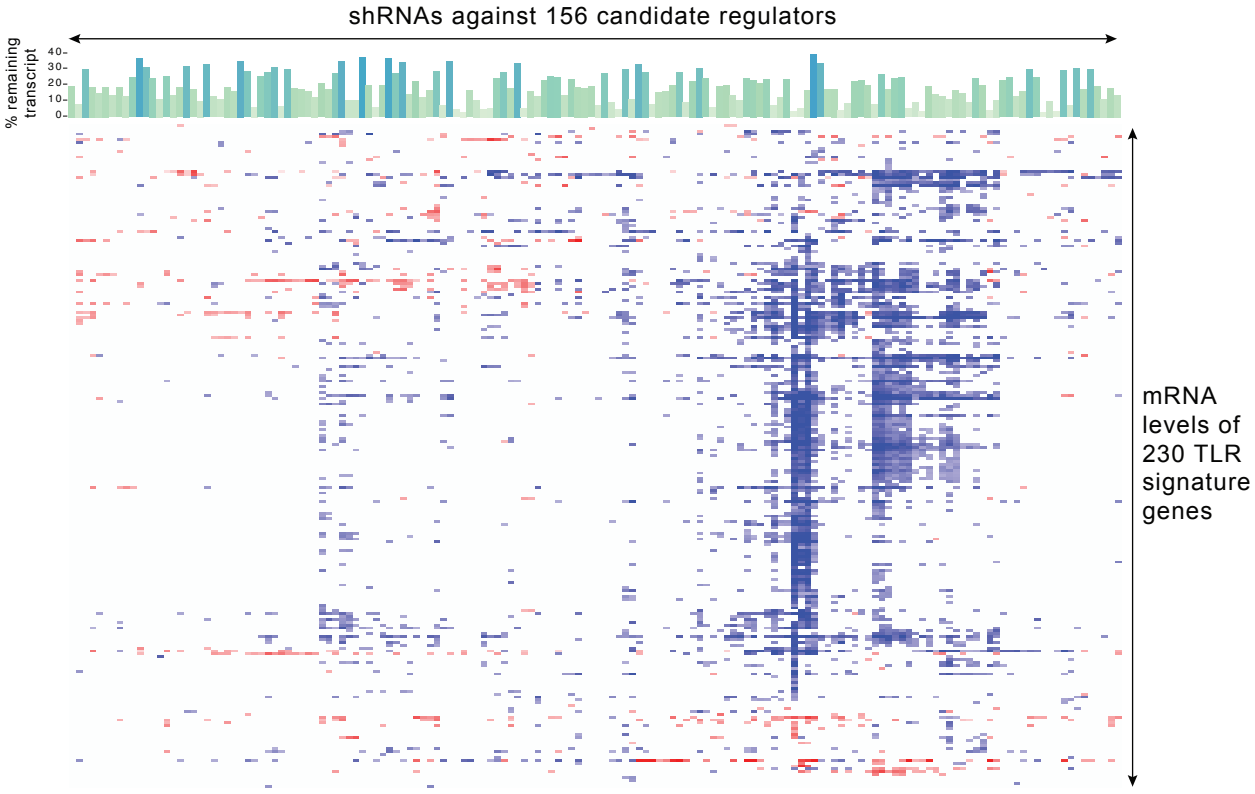

Figure 4.5. Genetic Perturbations of Candidate Regulators Affect TLR Signature Gene Expression
Shown are perturbation profiles of the 156 regulators with efficient shRNA-mediated knockdown (out of 203 initially selected), as indicated by % of remaining transcript compared to control shRNAs (bar graph on top). The heatmap shows each perturbed regulator (columns) and their statistically significant effects on each of 230 TLR signature genes (rows). Blue: target gene expression decreased following perturbation; red: target gene expression increased following perturbation; white: no significant effect. Rows and columns were hierarchically clustered (Pearson's correlation).

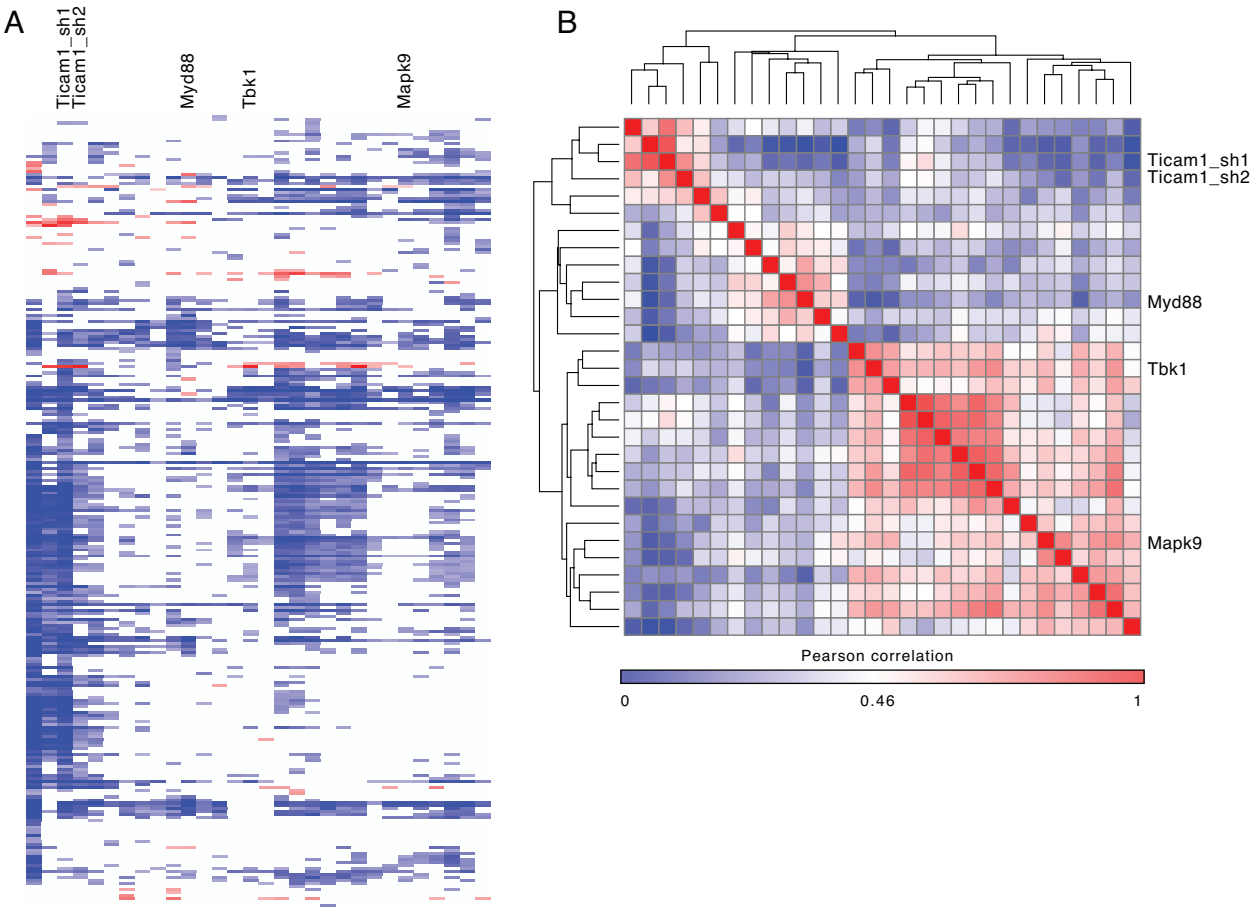

Figure 4.6. Functional Characterization Based on Similarity of Perturbation Profiles Reveals Genetic Interactions Among 30 Signaling Regulators

(A) Shown are perturbation profiles of the 30 regulators with significant effects on TLR signature genes (extracted from the matrix in Figure 4.5).
(B) Shown is a correlation matrix of the perturbation profiles from the 30 candidate regulators that significantly affected TLR signature gene expression (from **Figure 4.5**). Pearson's correlation was used to compute pairwise correlation coefficients among all 30 profiles. Red: maximum correlation; blue: negative correlation.

These observations raised two related questions: (1) what are the roles of these 30 regulators and how do they function together, and (2) what is the overall organization of the 'signaling-to-transcription' TLR network? The following sections 4.4 and 4.5 present preliminary data and analyses that start to address these questions.

4.5. Systematic Functional Characterization of 30 Candidate Regulators Affecting TLR Signature Gene Expression

Next, we sought to systematically characterize the function of these 30 candidate regulators. To this end, we designed and applied in DCs three complementary proteomics-based assays to: (1) connect a regulator with downstream phosphorylation events, (2) identify the substrates of kinases, and (3) place each component within its functional context by identifying its binding partners (**Figure 4.7**).

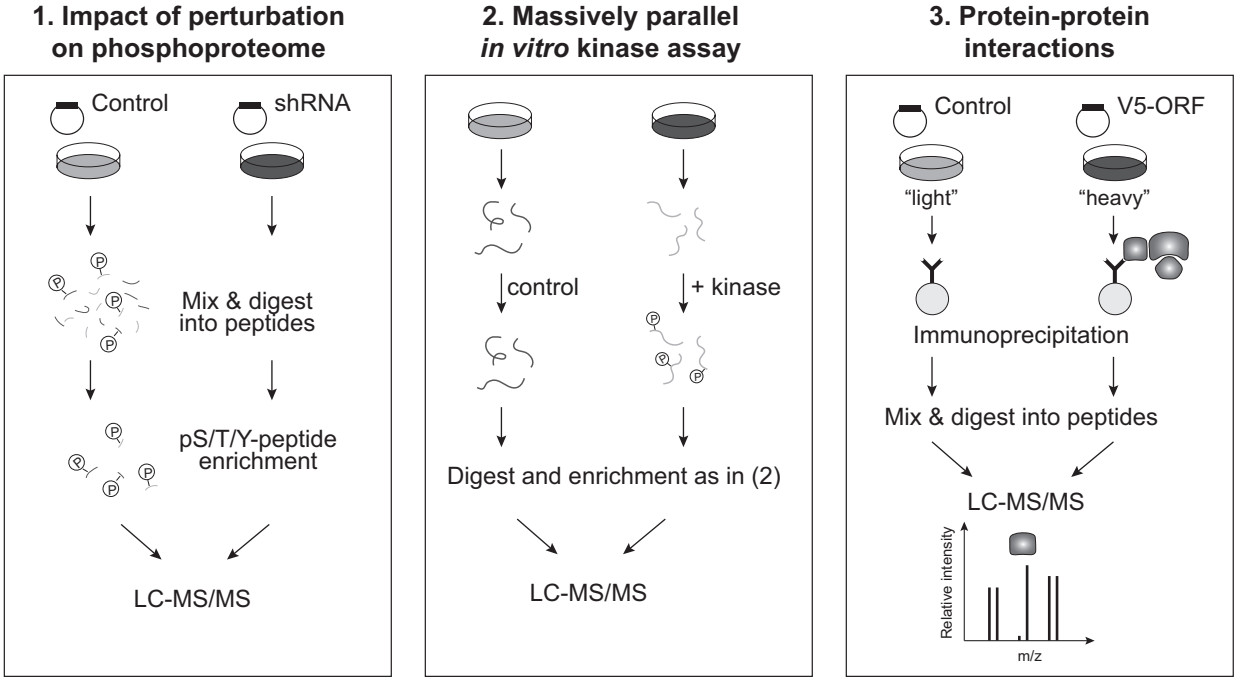

Figure 4.7. Functional Assays for Mechanistic Characterization of Candidate Regulators
Overview of the experimental workflow used for each large-scale biochemical assays described in the text on: (1) dissecting protein complexes, (2) measuring the impact of individual gene knockdown on the phosphoproteome, and (3) connecting kinases with their downstream substrates.

First, to connect each candidate regulators with downstream phosphorylation events (direct and indirect), we will genetically perturb each gene and measure the impact on the DC phosphoproteome upon LPS stimulation at a single time point (30 min) (**Figure 4.7, panel 1**). We and others have demonstrated in DCs (Chevrier et al., 2011), and in other mammalian or

yeast models (Bodenmiller et al., 2010; Grosstessner-Hain et al., 2011; Hsu et al., 2011; Koch et al., 2011; Santamaria et al., 2011), the validity of this approach in determining functional pathway components downstream of a given network node. Second, for the 10 protein kinases present among the 30 candidate regulators, we will determine their direct substrates using a massively parallel *in vitro* kinase assay, relying on purified recombinant kinases incubated with DC protein lysates followed by phoshoproteomic measurements (**Figure 4.7, panel 2**). Preliminary datasets using recombinant Tbk1 protein kinase demonstrated the feasibility of this method by identifying its known substrate motif (data not shown). Finally, we will place each component within their functional context in DCs by pulling them down, and identifying their binding partners upon stimulation (**Figure 4.7, panel 3**). The first two assays delineated above are part of ongoing efforts, and we will focus on the preliminary results obtained for the pull-down experiments.

To start investigating the precise molecular mechanisms of action of interesting signaling nodes or modules emanating from the results mentioned above, we used affinity purification followed by mass spectrometry (AP-MS). V5-tagged open reading frames (ORFs) were expressed in DCs using a lentiviral vector. For preliminary development of the methodology, we focused on Myd88, well-described proximal adaptor in the TLR4 signaling cascade (**Figure 4.8A**).

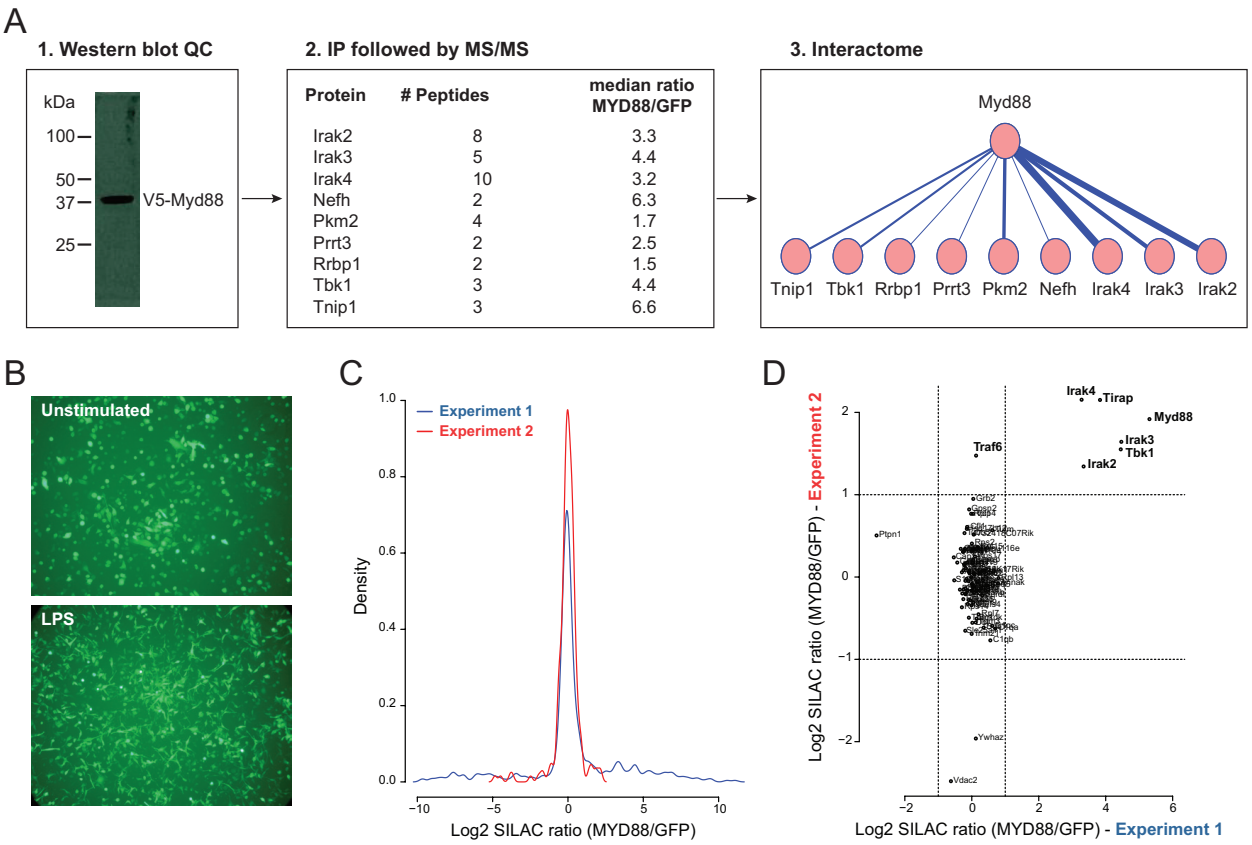

Figure 4.8. Myd88 Affinity-Purification Followed by Mass Spectrometry (AP-MS) Reveals Known and Candidate Interaction Partners

(A) Overview of the experimental workflow. V5-tagged Myd88 ORF was overexpressed in BMDCs using lentiviral delivery followed by blasticidin selection. Verification of ORF expression and size was done by Western blot using anti-V5 mAb (panel 1). For AP-MS experiments, cells were grown in SILAC media for quantification ("light" for GFP-V5, and "heavy" for Myd88-V5) and stimulated with LPS for 30 min before lysis and separate IPs (for each ORF-V5) using magnetic beads covalently linked to anti-V5 mAb. After washes, on-bead protein complexes from each IP (GFP-V5 and Myd88-V5) were mixed, digested into peptides, and analyzed by mass spectrometry.
(B) DCs overexpressing V5-tagged ORFs normally respond to LPS. BMDCs were infected with GFP-V5 lentivirus and selected with blasticidin. Shown are immunofluorescence images of GFP-V5-expressing BMDCs before (top) and after (bottom) LPS stimulation for 4 h.
(C-D) Independent AP-MS experiments identify high-confidence Myd88-associated proteins. Shown are density plots of the SILAC ratios (Myd88/GFP) of two independent experiments (C), and a plot of the SILAC ratios for proteins identified in both of these experiments (D, proteins in the top-right corner representing high-confidence Myd88-associated proteins).

DCs subjected to lentiviral ORF delivery and expression exhibited normal responsiveness to LPS stimulation (**Figure 4.8B**). Following immunoprecipitation of Myd88-V5 and GFP-V5 (as a control) in parallel in DCs stimulated with LPS for 30 min, we identified high-confidence binding partners that include known interactors (Tirap, Irak family) as well as potential new ones (Tbk1)

(**Figure 4.8C and 4.8D**). Importantly, immunoprecipitation of endogenous Myd88 in DCs led to comparable results, including the identification or Iraks and Tbk1 as Myd88 interaction partners (data not shown). Furthermore, immunoprecipitation of Irak2-V5 led to the identification of endogenous Myd88 as well as most of the other Myd88-associated proteins identified by Myd88-V5 AP-MS (data not shown). These preliminary datasets reinforce the validity of using our AP-MS approach to obtain mechanistic clues about the function of known and candidate TLR pathway components.

4.6. Integrative Analysis of Physical and Functional Interactions to Model 'Signaling-to-Transcription' Events in TLR Responses

Finally, to investigate the organization of the 'signaling-to-transcription' pathways controlling TLR responses, the orthogonal datasets detailed above will be integrated into a statistical model. The goal of this integrative analysis is two-fold. First, our data-driven model will help extracting the general organization of the overlapping pathways suggested by the genetic interactions observed in **Figure 4.6**. For example, it will inform whether these pathways cross talk via shared components or complexes, what are the sequence of events from ligand recognition to transcription initiation, or what are the hierarchical relationships between signaling components. Second, it will provide mechanistic clues regarding the roles of signaling regulators in the control of TLR-induced transcripts. For example, it will generate specific hypotheses about the mechanisms of action of a new component or module within the TLR network, or even directly pinpoint this mechanism in certain cases. Because this work is ongoing, we simply present a preliminary workflow for data integration, together with preliminary results on an example of integrating functional and physical datasets.

In a nutshell, our data-driven modeling approach relies on the integration of physical and functional interactions emanating from the measurements reported above (**Figure 4.9A**). Physical interactions (**Figure 4.9A**) comprise measurements of phosphorylation (**Figure 4.2**), protein binding (AP-MS; **Figure 4.7 and 4.8**), kinase substrates (**Figure 4.7**, panel 2), and transcription factor (TF) binding sites (Garber et al., unpublished). Functional interactions (**Figure 4.9A**) include the effects of genetic perturbations on the TLR-induced transcripts (**Figure 4.5**) and phosphoproteome (**Figure 4.7**, panel 1). To integrate these orthogonal datasets into a model, we will apply a stepwise procedure: stimulus-dependent phosphorylation events identify signaling components and provide clues on temporality (*i.e.*, order of events) (**Figure 4.9B**, panel 1); effects of genetic perturbations on TLR-induced transcripts establish functional relationship for a given component and between signaling components (**Figure 4.9B**, panel 2); effects of genetic perturbations on the TLR-induced phosphoproteome together with kinases' substrates connect modules of components acting in a sequential signaling path (**Figure 4.9B**, panel 3); and protein-protein and protein-DNA interactions will inform the physical relationships between components and downstream gene regulation (**Figure 4.9B**, panel 4). Furthermore, if proven necessary to expand or increase the confidence of the resulting model network, we will complement our observed interactions using a custom database of >50,000 physical interactions with experimental evidence retrieved from several public repositories (BioGrid, HPRD, PhosphoSite, SPIKE, and NetPath).

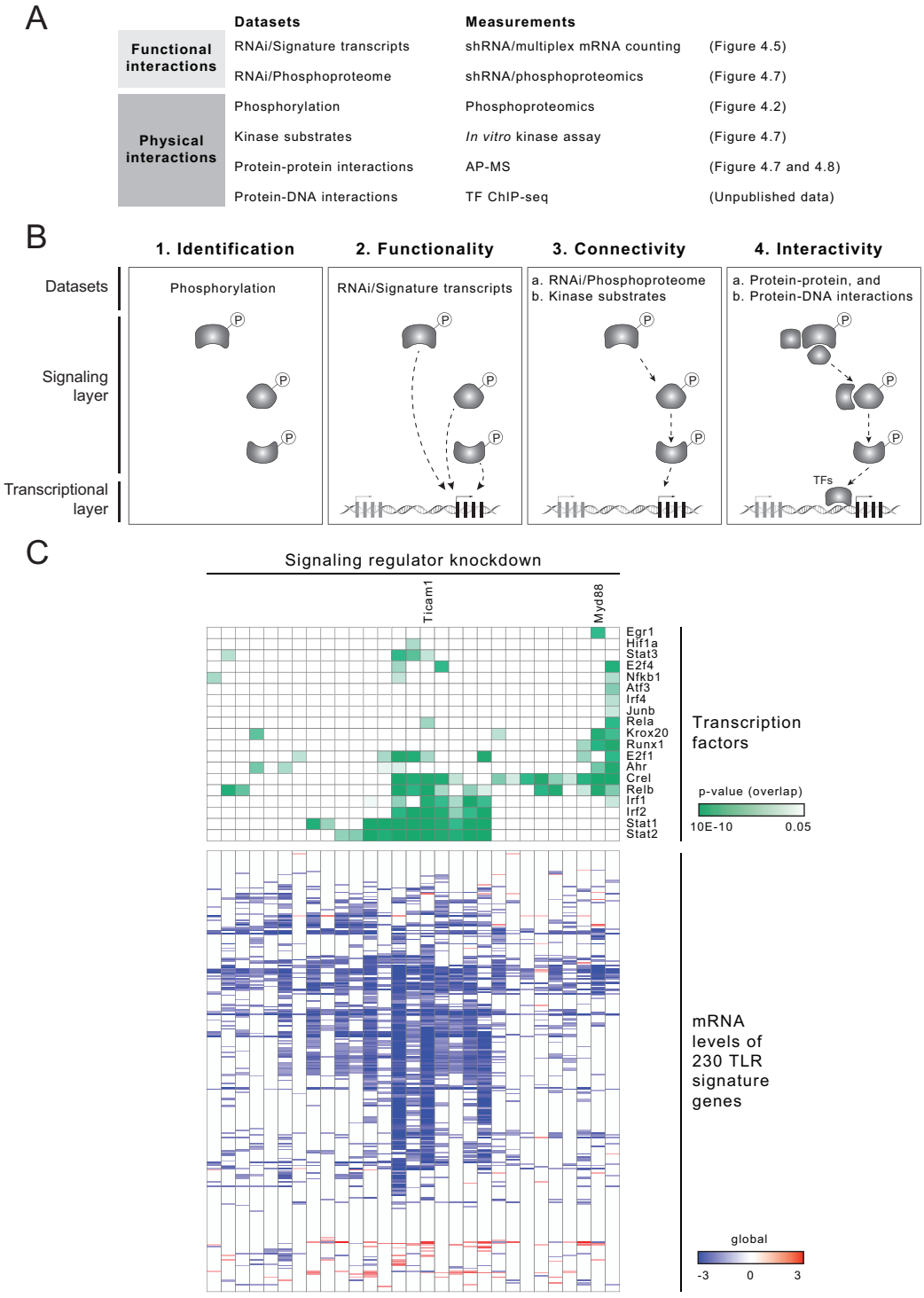

Figure 4.9. An Integrative Analysis to Dissect the Rules of 'Signaling-to-Transcription' Downstream of TLR4

(A) Schematic summary of the data types and measurements of this study. Datasets are categorized into functional and physical interactions, and corresponding measurement types and associated figures are listed on the right.

(B) Schematic overview of the proposed analytic framework to integrate functional and physical interactions into a model. From left to right: (1) Identification – candidate regulators are identified by

Figure 4.9 (Continued)
phosphoproteomic measurements (**Figure 4.2, 4.3, and 4.4**); (2) Functionality – regulators participating in the control of TLR signature genes are determined by functional genomics (**Figure 4.5**); (3) Connectivity – regulators are connected between themselves and with downstream effectors using functional phosphoproteomics and massively parallel *in vitro* kinase assay (**Figure 4.7**, panel 1 and 2); and (4) Interactivity – Physical interactions are determined for regulator-associated proteins (**Figure 4.8**) and for transcription factors (TFs) bound to promoter regions of TLR signature genes (Garber et al., unpublished data).
(C) An example of integrative analysis of functional and physical interactions to connect TLR signaling and transcriptional regulatory layers. Top matrix: shown are the 30 signaling regulators (columns) and TFs (rows) showing significant overlap (hypergeometric test) in terms of TLR signature genes affected by knockdown and bound at promoter regions, respectively. Green: significant correlation; white: no correlation. Bottom matrix: Shown are perturbation profiles of the signaling regulators from the top matrix. The heatmap shows each perturbed signaling regulator (columns, order identical as top matrix) and their statistically significant effects on each of 230 TLR signature genes (rows). Blue: target gene expression decreased following perturbation; red: target gene expression increased following perturbation; white: no significant effect.

As an initial example of the analyses described above, we show connections between the signaling and transcriptional layers of the TLR system. To connect these regulatory layers, we previously relied on genetic interactions inferred from similar perturbation profiles (**Chapter 3, section 3.3**). Here we go further by measuring overlaps (hypergeometric test) between the genes whose promoters are bound by 19 TFs of the TLR response (Garber et al., unpublished), and the genes whose transcripts were significantly impacted by knockdown of the 30 signaling regulators (**Figure 4.5**). By intersecting these physical and functional interactions, we identify significant overlaps highlighting that signaling regulators with similar perturbation profiles lie upstream of the same set of TFs (**Figure 4.9C**). For example, Myd88 clearly lies upstream of NF-κB family members Crel and Relb, and Ticam1-dependent genes significantly overlap with both inflammatory and antiviral TFs, such NF-κB and IRFs, respectively (**Figure 4.9C**). These observations are consistent with previous knowledge of the TLR pathways (Takeuchi and Akira, 2010). Further analyses following the preliminary framework delineated above will help to uncover the intermediate effectors and mechanistic events between the TLR signaling and transcriptional regulatory layers.

4.7. Discussion

By integrating functional and physical interactions about signaling and transcriptional regulatory layers, we uncovered a set of 30 TLR signaling regulators that control downstream transcriptional changes. Ongoing mechanistic characterization of these 30 regulators and integrative analysis of the resulting datasets are likely to help uncover modules and organizational rules driving the 'signaling-to-transcriptional' TLR network in DCs. Furthermore, the approach, datasets, and ongoing analyses reported here suggest multiple avenues for future exploration; we highlight below a handful of them.

The temporal phosphoproteomic datasets have been specifically mined for extraction of candidate signaling regulators. However, these highly information-rich data are likely to be useful for multiple other analytic, mining, and hypothesis-generating purposes. For example, intersecting phosphorylation data with other post-translational modifications (PTMs; *e.g.*, ubiquitinylation, acetylation) might help uncover regulatory codes on proteins harboring unique or combinations of these PTMs in response to LPS, as suggested by the p53 example (Sims and Reinberg, 2008). While we focuses on enriched enzymes and enzyme regulators likely to be active in signaling events in this work, our data uncovered multiple other classes of regulators that may be nominated as candidates such as TFs and other transcriptional regulators. Previous studies in macrophages (Weintz et al., 2010) or T cells (Navarro et al., 2011) have demonstrated that changes in phosphorylation may indeed be indicative of functional transcriptional regulators. Another potential avenue for future mining of these temporal phosphoproteomic data might come from searching for regulators of the positive and negative feedback regulatory loops involved in the TLR system (Rothlin et al., 2007).

Our perturbation analyses focused on measuring the effects of a given regulator knockdown on selected, downstream transcriptional targets. As sequencing costs keep decreasing, applying

the same approach to genome-wide readouts (*e.g.*, using RNA-seq) might help uncovering further modular subdivisions in the architecture of the signaling-to-transcription TLR circuits. For example, it might help further refine the various and overlapping branches of signaling events that culminate with changes in given sets of target genes.

Many candidate regulators did not have a significant impact on the TLR gene signature. Some of the possible explanations to this phenomenon include: (1) technical issues (*e.g.*, knockdown efficiency not sufficient), (2) redundancies (*e.g.*, PLK example from Chapter 3), and (3) potential importance of these genes not in controlling downstream transcription, but rather, other aspects of DC biology (*e.g.*, antigen processing and presentation, cell shape and motility, etc.), as exemplified in recent work by others (Neefjes et al., 2011; Paul et al., 2011). These results suggest the importance of conducting systematic analyses of epistatic interactions, as demonstrated in yeast for example (Collins et al., 2007; Roguev et al., 2008).

4.8. Experimental Procedures

Preparation of dendritic cells

Bone marrow-derived dendritic cells (BMDCs) were generated from 6-8 week old female C57BL/6J mice (Jackson Laboratories). Bone marrow cells were collected from femora and tibiae and plated at 10^6 cells/mL on non-tissue culture treated petri dishes in RPMI-1640 medium (Gibco), supplemented with 10% FBS, L-glutamine, penicillin/streptomycin, MEM non-essential amino acids, HEPES, sodium pyruvate, β-mercaptoethanol, and murine GM-CSF (15 ng/mL; Peprotech) or human Flt3L (100 ng/mL; Peprotech). GM-CSF-derived BMDCs were used directly for all RNAi experiments unsing lentiviral shRNAs. For all other experiments, floating cells from GM-CSF cultures were sorted at day 5 by MACS using the CD11c (N418) MicroBeads kit (Miltenyi Biotec). Sorted $CD11c^+$ cells were used as GM-CSF-derived BMDCs,

and plated at 10^6 cells/mL and stimulated at 16 h post sorting with 100 ng/mL ultra-pure *E. coli* K12 LPS (Invivogen).

Metabolic labeling of cells

For stable isotope labeling of amino acids in cell culture (SILAC) experiments, GM-CSF-derived BMDCs were grown for seven days in media containing either normal L-arginine (Arg-0) and L-lysine (Lys-0) (Sigma) or L-arginine $^{13}C_6$-$^{15}N_4$ (Arg-10) and L-lysine $^{13}C_6$-$^{15}N_2$ (Lys-8) (Sigma Isotec). Concentrations for L-arginine and L-lysine were 42 mg/L and 73 mg/L, respectively. To prevent metabolic conversion of L-arginine to L-proline we added 200 mg/L L-proline to the cell culture medium. The cell culture media, Roswell Park Memorial Institute-1640 (RPMI) deficient in L-arginine and L-lysine, was a custom media preparation from Caisson Laboratories (North Logan, UT) and dialyzed serum was obtained from SAFC-Sigma. We followed all standard SILAC media preparation and labeling steps as previously described (Ong and Mann, 2006).

Global serine, threonine, and tyrosine phosphorylation analysis

BMDCs grown in SILAC media were stimulated with LPS, and lysed and processed for enrichment of phosphopeptides as described in full in Chapter 3 (section 3.12).

Analysis of relative protein expression

BMDCs grown in SILAC media were left untreated or stimulated with LPS for 2 and 6 h. After in solution digest and desalting, 100 µg of total peptides were separated using an Agilent 3100 Offgel fractionator (Agilent, G3100A) as described in the manual. For separation into 12 fractions, we used Immobiline DryStrips, 13cm, pH 3-10 (GE Healtcare, 17-6001-14) that were rehydrated in a 1:50 dilution of IPG buffer, pH 3-10 (GE Healthcare, 17-6000-87) containing 5% glycerol. Peptides were reconstituted in IPG buffer (1:50 dilution) containing 5% glycerol and focused for 20kV*h with a maximum current of 50 µA and power of 200 mW. After separation,

fractions were acidified by adding 1% formic acid and desalted using StageTips. Peptide samples were analyzed on a LTQ Orbitrap Velos instrument.

Analysis of phosphoproteomic data

For supervised *k*-means clustering to partition the differentially expressed phosphopeptides, we set the parameter k (number of clusters) as the minimal number of clusters that provided a sufficient level of within-cluster similarity. For every cluster we define the within-cluster similarity as the average r^2 between the members of the cluster and the centroid of the cluster. We used the following cutoffs for: the minimum within-cluster similarity (across all clusters) to be >0.7, and the average (across all clusters) to be >0.75.

Pathway enrichment analyses were performed using the following databases and online resources: Panther (http://www.pantherdb.org), DAVID (http://david.abcc.ncifcrf.gov), MSigDB (http://www.broadinstitute.org/gsea/msigdb). Data were displayed using the Enrichment Map plugin (Merico et al., 2011) in Cytoscape (http://www.cytoscape.org/), and the software Gene-E (http://www.broadinstitute.org/cancer/software/GENE-E/)

shRNA knockdowns

High titer lentiviruses encoding shRNAs targeting genes of interest were obtained from The RNAi Consortium (TRC; Broad Institute). Bone marrow cells were infected with lentiviruses and processed for RNA extraction and measurments as described in Chapter 3 (section 3.12).

mRNA counting and data analysis

Samples were processed as described in Chapter 3 (section 3.12). We used a custom CodeSet constructed to detect a total of 267 genes (including 16 control genes whose expression remain unaffected by TLR stimulation) selected using the *GeneSelector* algorithm and manually populated for DC- or TLR-related genes of interest.

Multiplex mRNA count data analysis

For normalization, we divide the nCounter (NanoString) mRNA count values for each gene by the sum of counts obtained for the 16 control genes present in our custom nCounter CodeSet. To determine significantly affected signature genes, a fold-change ratio is computed for each pairwise comparison of a knockdown sample versus a set of control samples (*i.e.*, non-targeting shRNA; at least 10 per experimental batch). As a threshold, we require a substantial fold-change (above a threshold value t) in the same direction (up- or down-regulation) in more than half of the pairwise comparisons sample vs. control shRNA. The threshold value t is determined as max (q, d), d being the mean + 1.645 times the standard deviation in the fold change shown by the control genes (corresponding to p = 0.05, under the assumption of normality). The threshold q is similar for all comparisons and is based on the noise level estimated from the control shRNA samples. Specifically, we compute gene expression fold changes in all possible pairs of control shRNA samples (which are supposed to be consistent). We set the threshold q such that 95% of the comparisons exhibit lower fold change than q. The resulting value of q is 1.961. Notably, we ignore all pairwise comparisons in which both control and knockdown samples had low counts before normalization (<50). All heatmaps and distance matrix analyses were generated using the software Gene-E, referenced above.

Affinity purification followed by mass spectrometry

Analysis of interaction partners of V5-tagged proteins was performed using a fast and low-stringency single-step purification procedure (to retain weak binders and potentially transient interactions) previously decribed by Matthias Mann's group (Hubner et al., 2010; Hubner and Mann, 2011), with several modifications to fit our experimental system. 2 x 10^6 bone marrow cells were plated in SILAC complete medium supplemented with 15 ng/mL GM-CSF in 10-cm Petri dishes, and infected two days later with lentiviruses (MOI ~10-20) containing V5-tagged

ORFs (Yang et al., 2011) in 10-cm Petri dishes. 2-4 h after infection, cells were fed with GM-CSF-containing complete medium. Two days post-infection, GM-CSF-containing complete medium supplemented with blasticidin (10 µg/mL) was added to cells, which were further incubated for 3 days. ORF expression and size was validated using standard Western blotting with anti-V5 antibody (Invitrogen).

For immunoprecipitation (IP) of protein complexes, BMDCs expressing a V5-tagged ORF encoding MYD88 were stimulated with LPS for 30 min, scraped on ice and washed in ice-cold PBS. Cell pellets were lysed for 30 min on ice in a lysis buffer containing 150 mM NaCl, 50 mM Tris pH 7.5, 5% Glycerol, 1% IGPAL-CA-630 (Sigma, #I8896), and freshly added protease and phosphatase inhibitors (Roche). After centrifugation at 4°C for 10 min at 14,000 g, protein concentration in supernatants was measured by BCA (Pierce), and equal amounts (~2.5-3 mg) of lysates from each SILAC samples were used for subsequent IP. Cell lysates were incubated for ~16 h at 4°C on a roller with anti-V5 tag antibody covalently bound to magnetic beads (MBL). Beads were then washed twice with wash buffer (150 mM NaCl, 50 mM Tris pH 7.5, 5% Glycerol) containing 0.05% 1% IGPAL-CA-630, and twice with wash buffer alone. Beads from each SILAC state were combined after the first wash. Purified protein complexes were then eluted non-specifically by direct on-bead digestion with trypsin using a buffer containing 2 M urea, 50 mM Tris pH 7.5, 1 mM DTT, and 5 µg/mL Trypsin. After elution, samples were reduced (4 mM DTT) and alkylated (10 mM iodoacetamide) following standard procedures, and further digested with trypsin overnight. Digestion were stopped by adding 1% TFA, and peptides were desalted purified on C_{18} StageTips.

Desalted peptide samples were separated on an online nanoflow UHPLC system (Proxeon EASY-nLC 1000) and analyzed on a Q Exactive (Thermo Fisher Scientific) mass spectrometer. We used a 13 cm reversed phase fused-silica capillary column (New Objective, PicoFrit PF360-75-10-N-5 with 10 µm tip opening and 75 µm inner diameter) packed in-house with 3 µm ReproSil-Pur C18-AQ media (Dr. Maisch GmbH) and separated peptides at a flow rate of 200

nL/min in a 82 min linear gradient from 6 to 30% composition of solvent A (3% acetonitrile /0.1% formic acid) and solvent B (90% acetonitrile /0.1% formic acid). The Q Exactive was operated at a spray voltage of 2 kV, a capillary temperature of 250 C and a S-lens RF level of 50. Data was acquired in positive ion mode, with MS1 scans at a resolution of 70,000 at m/z=200, a mass range from 300-1800, AGC target of 1e6 and 5 ms maximum ion time. Up to 12 of the most intense ions per duty cycle were isolated using an isolation window of 2.5 m/z and fragmented by HCD at a NCE of 25 with an underfill ratio set at 5%. For data dependent MS2 scans we used a resolution of 17,500, an AGC target of 5e4 and a maximum ion time of 120ms. All ions selected for MS2 scans were dynamically excluded for 20 s after fragmentation.

4.9. References

Barton, G.M., and Kagan, J.C. (2009). A cell biological view of Toll-like receptor function: regulation through compartmentalization. Nature reviews Immunology *9*, 535-542.

Bodenmiller, B., Wanka, S., Kraft, C., Urban, J., Campbell, D., Pedrioli, P.G., Gerrits, B., Picotti, P., Lam, H., Vitek, O.*, et al.* (2010). Phosphoproteomic analysis reveals interconnected system-wide responses to perturbations of kinases and phosphatases in yeast. Science signaling *3*, rs4.

Chevrier, N., Mertins, P., Artyomov, M.N., Shalek, A.K., Iannacone, M., Ciaccio, M.F., Gat-Viks, I., Tonti, E., DeGrace, M.M., Clauser, K.R.*, et al.* (2011). Systematic discovery of TLR signaling components delineates viral-sensing circuits. Cell *147*, 853-867.

Choudhary, C., and Mann, M. (2010). Decoding signalling networks by mass spectrometry-based proteomics. Nature reviews Molecular cell biology *11*, 427-439.

Collins, S.R., Miller, K.M., Maas, N.L., Roguev, A., Fillingham, J., Chu, C.S., Schuldiner, M., Gebbia, M., Recht, J., Shales, M.*, et al.* (2007). Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. Nature *446*, 806-810.

Grosstessner-Hain, K., Hegemann, B., Novatchkova, M., Rameseder, J., Joughin, B.A., Hudecz, O., Roitinger, E., Pichler, P., Kraut, N., Yaffe, M.B.*, et al.* (2011). Quantitative phospho-proteomics to investigate the polo-like kinase 1-dependent phospho-proteome. Molecular & cellular proteomics : MCP *10*, M111 008540.

Hsu, P.P., Kang, S.A., Rameseder, J., Zhang, Y., Ottina, K.A., Lim, D., Peterson, T.R., Choi, Y., Gray, N.S., Yaffe, M.B.*, et al.* (2011). The mTOR-regulated phosphoproteome reveals a mechanism of mTORC1-mediated inhibition of growth factor signaling. Science *332*, 1317-1322.

Hubner, N.C., Bird, A.W., Cox, J., Splettstoesser, B., Bandilla, P., Poser, I., Hyman, A., and Mann, M. (2010). Quantitative proteomics combined with BAC TransgeneOmics reveals in vivo protein interactions. The Journal of cell biology *189*, 739-754.

Hubner, N.C., and Mann, M. (2011). Extracting gene function from protein-protein interactions using Quantitative BAC InteraCtomics (QUBIC). Methods *53*, 453-459.

Koch, A., Krug, K., Pengelley, S., Macek, B., and Hauf, S. (2011). Mitotic substrates of the kinase aurora with roles in chromatin regulation identified through quantitative phosphoproteomics of fission yeast. Science signaling *4*, rs6.

Lee, C.C., Avalos, A.M., and Ploegh, H.L. (2012). Accessory molecules for Toll-like receptors and their function. Nature reviews Immunology *12*, 168-179.

Merico, D., Isserlin, R., and Bader, G.D. (2011). Visualizing gene-set enrichment results using the Cytoscape plug-in enrichment map. Methods Mol Biol *781*, 257-277.

Navarro, M.N., Goebel, J., Feijoo-Carnero, C., Morrice, N., and Cantrell, D.A. (2011). Phosphoproteomic analysis reveals an intrinsic pathway for the regulation of histone deacetylase 7 that controls the function of cytotoxic T lymphocytes. Nature immunology *12*, 352-361.

Neefjes, J., Jongsma, M.L., Paul, P., and Bakke, O. (2011). Towards a systems understanding of MHC class I and MHC class II antigen presentation. Nature reviews Immunology *11*, 823-836.

Ong, S.E., and Mann, M. (2006). A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). Nature protocols *1*, 2650-2660.

Paul, P., van den Hoorn, T., Jongsma, M.L., Bakker, M.J., Hengeveld, R., Janssen, L., Cresswell, P., Egan, D.A., van Ham, M., Ten Brinke, A.*, et al.* (2011). A Genome-wide multidimensional RNAi screen reveals pathways controlling MHC class II antigen presentation. Cell *145*, 268-283.

Roguev, A., Bandyopadhyay, S., Zofall, M., Zhang, K., Fischer, T., Collins, S.R., Qu, H., Shales, M., Park, H.O., Hayles, J.*, et al.* (2008). Conservation and rewiring of functional modules revealed by an epistasis map in fission yeast. Science *322*, 405-410.

Rothlin, C.V., Ghosh, S., Zuniga, E.I., Oldstone, M.B., and Lemke, G. (2007). TAM receptors are pleiotropic inhibitors of the innate immune response. Cell *131*, 1124-1136.

Santamaria, A., Wang, B., Elowe, S., Malik, R., Zhang, F., Bauer, M., Schmidt, A., Sillje, H.H., Korner, R., and Nigg, E.A. (2011). The Plk1-dependent phosphoproteome of the early mitotic spindle. Molecular & cellular proteomics : MCP *10*, M110 004457.

Sims, R.J., 3rd, and Reinberg, D. (2008). Is there a code embedded in proteins that is based on post-translational modifications? Nature reviews Molecular cell biology *9*, 815-820.

Takeuchi, O., and Akira, S. (2010). Pattern recognition receptors and inflammation. Cell *140*, 805-820.

Weintz, G., Olsen, J.V., Fruhauf, K., Niedzielska, M., Amit, I., Jantsch, J., Mages, J., Frech, C., Dolken, L., Mann, M.*, et al.* (2010). The phosphoproteome of toll-like receptor-activated macrophages. Molecular systems biology *6*, 371.

Yang, X., Boehm, J.S., Yang, X., Salehi-Ashtiani, K., Hao, T., Shen, Y., Lubonja, R., Thomas, S.R., Alkan, O., Bhimdi, T.*, et al.* (2011). A public genome-scale lentiviral expression library of human ORFs. Nature methods *8*, 659-661.

Chapter 5 – Concluding Remarks

The approaches and results presented here set the ground for further developments from both systems biological and immunological standpoints. For example, immediate applications of these approaches can be foreseen with the use of additional data sources for candidate regulator nomination (*e.g.*, genes emanating from genome-wide association studies). Furthermore, the study of additional regulatory layers embedded within the TLR system (*e.g.*, see Appendix A) will be key to reach a more detailed understanding of how it operates, and to potentially gain insights into basic cell biology. Equally important will be the integration of these regulatory layers with variables such as context (*e.g.*, diseases) and space (*e.g.*, intracellular structures). Furthermore, while the studies presented here focused on the TLR system, it is clear that many inter-dependencies exist among pathogen-sensing pathways (*e.g.*, interplay between TLRs and NLRs). Conducting systematic analyses of these interactions to identify their levels of interface and functional relevance will be key to crack the code of innate immune sensing and activation. Finally, a deeper understanding of pathogen-sensing pathways will not be reached unless placed in the context of evolutionary constraints and organism-level immune responses.

The advent of systems approaches has propelled biology in an era of accelerated discovery. Furthermore, a quantitative understanding of biological systems at a high level of resolution is likely to be within reach in the not too distant future. Some of the challenges for the next decade are to fully realize the vision and promise of systems approaches by leveraging their potential to catalyze the generation of refined hypotheses and precise cellular networks about the mechanisms of health and disease. Resulting knowledge will undoubtedly help (*i*) in manipulating biological systems either by rationale targeting of components or pathways, alone or in combination, and (*ii*) in building biological systems *de novo*.

Appendix A – Large Non-Coding RNAs as a Putative Regulatory Layer of the Toll-Like Receptor System

A.1. Introduction

In this appendix, we provide preliminary evidence that large non-coding RNAs (ncRNAs) might represent a previously unrecognized layer of regulation in the TLR system. The idea that RNA molecules would be well adapted for regulatory roles emerged more than 50 years ago (Eddy, 2001; Jacob and Monod, 1961). Initial discovery of functional RNAs included small nuclear (sn) and nucleolar (sno) RNAs, involved in ribosome biogenesis, and ribosomal RNAs and transfer RNAs. A few decades later, the first examples of functional, mRNA-like, large ncRNAs were discovered. While they resembled mRNAs in terms of length and structure (*i.e.*, capped, exonic), these large ncRNAs did not possess any protein-coding potential. For example, Xist has been shown to function in imprinting and X chromosome inactivation (Brown et al., 1991; Penny et al., 1996). Another example of functional, large ncRNA comes from T cell biology: NRON (noncoding repressor of NFAT) prevents NFAT from transclocating into the nucleus, therefore blocking NFAT-induced gene expression (Willingham et al., 2005).

Large-scale studies have focused on better characterizing eukaryotic transcripts present in a given cell or tissue under multiple conditions. Using massive shotgun sequencing of mouse complementary DNAs (cDNAs), a consortium study highlighted a broader transcriptional activity than anticipated (Carninci et al., 2005). Another approach relied on the development of high-density tiling arrays. These arrays were designed to represent genomic DNA at various sequence resolutions regardless of known gene annotations, thus providing a powerful and unbiased tool to investigate transcribed sequences of genomic chromosomes or loci (Bertone et

al., 2005). Through the interrogation of tiling arrays, thousands of new transcribed regions of the human genome were identified (Bertone et al., 2004). More recently, the combination of direct high-throughput sequencing of cDNAs and the use of tiling arrays allowed for the detection of widespread transcriptional activity in more than 90% of the genome of the yeast *Schizosaccharomyces pombe* (Wilhelm et al., 2008). Altogether, these datasets and those of others (Birney et al., 2007) demonstrated that eukaryotic genomes are almost entirely transcribed, producing multiple small (<200 nucleotides) and large ncRNAs (>200 nucleotides) transcripts. An important question arising from these datasets is whether these large ncRNAs represent simple transcriptional noise, or are part of a larger set of RNAs with regulatory roles in cells, similarly to the few examples of functional large ncRNAs mentioned above (Ponting et al., 2009).

A.2. Chromatin Maps Identify Large Intergenic Non-Coding RNAs

A recent study brought experimental evidence of the existence of a class of long ncRNAs, referred to here as long intergenic ncRNAs (lincRNAs) (Guttman et al., 2009). This work reports the identification of ~1600 lincRNAs in the mouse genome based on genome-wide modified-chromatin immunoprecipitation (ChIP) followed by massive parallel sequencing (ChIP-seq) in four cell types. This experimental approach rests on the observation that genes actively transcribed by polymerase (Pol) II are marked by Histone 3-Lysine 4 trimethylation (H3K4me3) at their promoter, and Histone 3-Lysine 36 trimethylation (H3K36me3) along the length of the transcribed region (Mikkelsen et al., 2007). These modified-chromatin domains (K4/K36 domains) are found in both protein-coding genes and well-known ncRNAs such as miRNAs. Here we asked whether these lincRNAs might play a role in regulating TLR responses of DCs.

A.3. A Set of LincRNAs is Differentially Regulated upon TLR Stimulation in DCs

First, we designed custom tiling arrays with probes tiling across the K4/K36 domains of recently identified lincRNAs and the 3' regions of ~300 protein-coding genes, which are differentially regulated in TLR-activated BMDCs (Amit et al., 2009). Both protein-coding genes known to be involved in the TLR pathways (data not shown) and a set of 24 lincRNAs followed three similar patterns of induction (**Figure A.1A**). Interestingly, some of our lincRNA candidates lie in the genome in the neighboring regions of IRF family members (**Figure A.1A**). Moreover, some of the promoter regions of these lincRNAs have binding sites for IRF2, IRF3 and STAT6 transcription factors (data not shown). However, only a small number of these lincRNAs were reproducible by qRT-PCR, which is likely due to: (1) the use of tiling array probes for lincRNA exonic regions identified in non-immune cells and that may not be relevant to DCs, and (2) the intrinsic limitations of tiling array platforms (*e.g.*, poor reproducibility, high background, cross-hybridization).

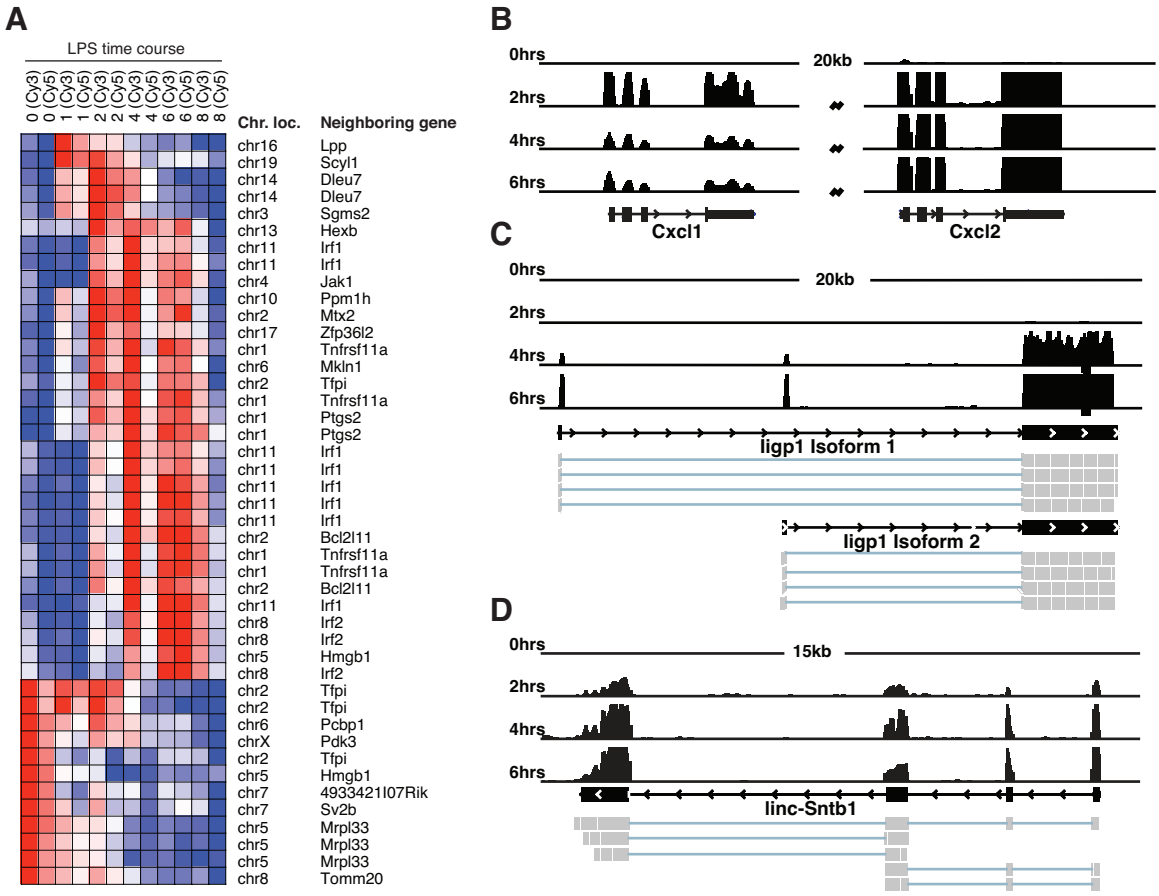

Figure A.1. Identification of DC-specific large intergenic noncoding RNAs (lincRNAs)

(A) Heatmap of lincRNA exons differentially regulated in DCs stimulated with LPS for indicated times. cDNA was hybridized on custom tiling arrays with probes for both lincRNA and protein-coding gene exons. Groups of two contiguous columns represent duplicates processed independently with dye swap (Cy3 or Cy5). On the right are indicated chromosomal location (Chr. loc.), and closest neighboring gene (downstream) for each lincRNA exon.

(B-D) RNA-seq of cDNA from BMDCs stimulated for 0, 2, 4, and 6 h with LPS. Massively parallel sequencing (Illumina platform) was performed on poly(A)+ mRNA, with 76-base paired-ends. Data analysis and transcriptome reconstruction were performed using the Scripture algorithm. Shown are expression levels of two well-known TLR-inducible genes (Cxcl1 and Cxcl2) as measured by number of reads (B), splicing information on Iigp1, an LPS-inducible gene (C), and an example of a differentially regulated lincRNA in DCs (*i.e.*, not expressed in MEFs, MLFs, NPCs, or ESCs) (D).

To address these problems, we moved to a different approach using RNA-seq. RNA-seq technologies have recently emerged as an unbiased way to precisely identify, quantify, and map the transcriptome of a cell (Wang et al., 2009). We generated poly(A)$^+$ RNA from a time course of LPS stimulated DCs, and sequenced 50 million 76-base paired-end reads on the Illumina GAII platform. Analyzing this data, we were able to identify known protein-coding genes (**Figure**

A.1B), their isoforms (**Figure A.1C**), as well as a set of DC-specific lincRNAs (**Figure A.1D**). Quantification based on RNA-seq read numbers closely mirrored what we previously observed for protein-coding genes using microarrays or qRT-PCR (data not shown).

We are now finalizing a list of ~25 most differentially regulated lincRNAs that are specific to DCs (i.e., not expressed in other mouse cell types such as ESCs, MEFs, NPCs, MLFs). Using RNA-seq data for the study of lincRNAs, or other ncRNA species in general, has two major advantages: (1) it gives greater confidence in calling true positive hits in terms of differential expression upon stimulation, and (2) it precisely delimits intron-exon boundaries as well as potential splice variants. Having this dataset at hand will greatly facilitate experimental queries of the functions of lincRNAs in the TLR system (e.g., shRNA design for knockdowns, cloning for overexpression).

Among the lincRNAs identified in these studies, we found linc-Ptgs2 next to the Cox-2 locus (**Figure A.2A**). Interestingly, it shares a similar pattern of induction as Cox-2 mRNA upon LPS stimulation in BMDCs (**Figure A.2B**). The differential expression between linc-Ptgs2 and Cox-2 across mouse tissues suggests that these 2 loci do not entirely share the same transcriptional regulators (**Figure A.2C**). Nonetheless both Cox-2 and linc-Ptgs2 are under the control of NF-κB in DCs. Indeed, knockdown of individual NF-κB family members leads to a decrease in linc-Ptgs2 and Cox2 transcripts upon perturbation of Nfkb1, Rela, and Relb, but not Rel and Nfkb2 (**Figure A.2D**). RNA fluorescence in situ hybridization (FISH) localized linc-Ptgs2 to the nucleus of LPS-treated BMDCs (**Figure A.2E**). Whether linc-Ptgs2 and the other DC-specific lincRNAs are functional, and if so, what are their mechanisms of action, remains to be determined.

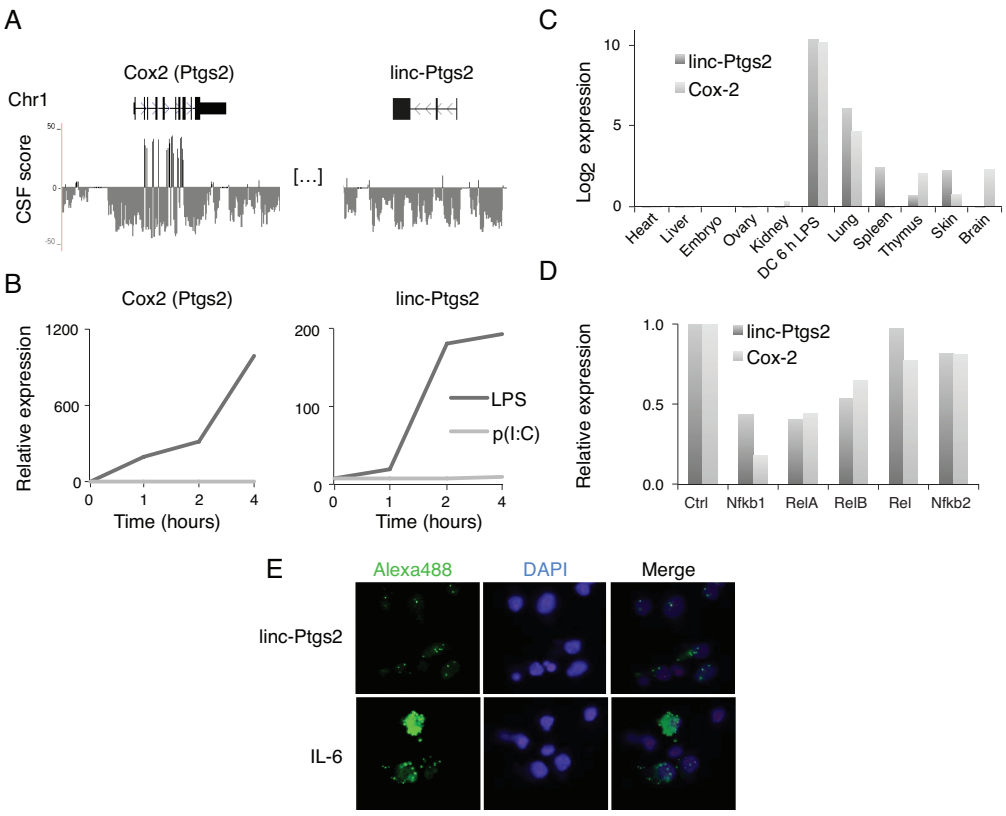

Figure A.2. linc-Ptgs2 is regulated by NF-κB family and preferentially localize to the nucleus

(A) linc-Ptgs2 does not exhibit any coding potential, as opposed to its closest neighboring gene *Ptgs2*. Shown are the codon substitution frequency (CSF) score, a positive score being indicative of protein-coding potential whereas a negative one indicates a lack thereof, for the *Ptgs2* and *linc-Ptgs2* loci.
(B) linc-Ptgs2 and Ptgs2 mRNA follow a similar pattern of induction following LPS, but not p(I:C). Shown are RNA expression levels (qPCR; normalized to Gapdh mRNA and relative to t = 0) of linc-Ptgs2 and Ptgs2 in DCs upon stimulation with LPS at indicated time points.
(C) linc-Ptgs2 expression pattern across mouse tissues. Shown are linc-Ptgs2 RNA levels in indicated tissues (qPCR; normalized to Gapdh mRNA).
(D) Identical NF-κB family members regulate both Ptgs2 and linc-Ptgs2 induction. Shown are RNA expression levels (qPCR; normalized to Gapdh mRNA) in DCs stimulated for 6 h with LPS in which indicated NF-κB members have been knocked down using shRNAs.
(E) linc-Ptgs2 preferentially localized to the nucleus. Shown are micrographs of RNA fluorescence in situ hybridizations (FISH) with probes for linc-Ptgs2 (top) and IL-6 mRNA (bottom) in LPS-treated DCs.

A.4. Future Prospects

Recent reports suggest a role for lincRNAs in regulating gene expression (Guttman and Rinn, 2012). Future studies will likely focus on some of the following questions and hypotheses: (*1*) how to build a comprehensive catalogue of lincRNAs and other small and large ncRNAs encoded by mammalian genomes, and how to categorize them based on functional properties

and/or structural similarities; (*2*) what are the features characterizing protein-lincRNA interactions (*e.g.*, sequence and/or structural motifs); (*3*) if lincRNAs modify the functional properties of their protein binding partner, how do they achieve doing so (*e.g.*, allostery); and (*4*) what are the subcellular (*e.g.*, nucleus vs. cytosol) and suborganellar (*e.g.*, specific nuclear locations) localizations characterizing lincRNAs and their biological roles?

A.5. References

Amit, I., Garber, M., Chevrier, N., Leite, A.P., Donner, Y., Eisenhaure, T., Guttman, M., Grenier, J.K., Li, W., Zuk, O.*, et al.* (2009). Unbiased reconstruction of a mammalian transcriptional network mediating pathogen responses. Science *326*, 257-263.

Bertone, P., Gerstein, M., and Snyder, M. (2005). Applications of DNA tiling arrays to experimental genome annotation and regulatory pathway discovery. Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology *13*, 259-274.

Bertone, P., Stolc, V., Royce, T.E., Rozowsky, J.S., Urban, A.E., Zhu, X., Rinn, J.L., Tongprasit, W., Samanta, M., Weissman, S.*, et al.* (2004). Global identification of human transcribed sequences with genome tiling arrays. Science *306*, 2242-2246.

Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigo, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., Thurman, R.E.*, et al.* (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature *447*, 799-816.

Brown, C.J., Ballabio, A., Rupert, J.L., Lafreniere, R.G., Grompe, M., Tonlorenzi, R., and Willard, H.F. (1991). A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. Nature *349*, 38-44.

Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C.*, et al.* (2005). The transcriptional landscape of the mammalian genome. Science *309*, 1559-1563.

Eddy, S.R. (2001). Non-coding RNA genes and the modern RNA world. Nature reviews Genetics *2*, 919-929.

Guttman, M., Amit, I., Garber, M., French, C., Lin, M.F., Feldser, D., Huarte, M., Zuk, O., Carey, B.W., Cassady, J.P.*, et al.* (2009). Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature *458*, 223-227.

Guttman, M., and Rinn, J.L. (2012). Modular regulatory principles of large non-coding RNAs. Nature *482*, 339-346.

Jacob, F., and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. Journal of molecular biology *3*, 318-356.

Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P.*, et al.* (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553-560.

Penny, G.D., Kay, G.F., Sheardown, S.A., Rastan, S., and Brockdorff, N. (1996). Requirement for Xist in X chromosome inactivation. Nature *379*, 131-137.

Ponting, C.P., Oliver, P.L., and Reik, W. (2009). Evolution and functions of long noncoding RNAs. Cell *136*, 629-641.

Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. Nature reviews Genetics *10*, 57-63.

Wilhelm, B.T., Marguerat, S., Watt, S., Schubert, F., Wood, V., Goodhead, I., Penkett, C.J., Rogers, J., and Bahler, J. (2008). Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. Nature *453*, 1239-1243.

Willingham, A.T., Orth, A.P., Batalov, S., Peters, E.C., Wen, B.G., Aza-Blanc, P., Hogenesch, J.B., and Schultz, P.G. (2005). A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. Science *309*, 1570-1573.

Appendix B – Publications from Chapter 2 and 3

Unbiased Reconstruction of a Mammalian Transcriptional Network Mediating Pathogen Responses

Ido Amit *et al.*

Science **326**, 257 (2009);
DOI: 10.1126/science.1179050

19. E. Y. Huang *et al.*, *J. Bacteriol.* **179**, 5648 (1997).
20. Y. A. Hannun, C. Luberto, *Trends Cell Biol.* **10**, 73 (2000).
21. P. W. van der Wielen *et al.*, *Science* **307**, 121 (2005).
22. M. M. Yakimov *et al.*, *Curr. Opin. Biotechnol.* **18**, 257 (2007).
23. A. F. Pronk *et al.*, *J. Bacteriol.* **177**, 75 (1995).
24. O. V. Golyshina, K. N. Timmis, *Environ. Microbiol.* **7**, 1277 (2005).
25. D. A. Pearce, F. Sherman, *J. Bacteriol.* **181**, 4774 (1999).
26. J. D. Woodson *et al.*, *J. Bacteriol.* **185**, 7193 (2003).
27. Y. Jiao *et al.*, *Appl. Environ. Microbiol.* **71**, 4487 (2005).
28. This research was supported by the BIO2006-11738, CSD2007-00005, GEN2006-27750-C-4-E, BFU2008-04398-E/BMC, and KBBE-226977 projects. A.B and Y.A-R thank the Spanish MEC for the FPU and FPI fellowships. F.P. thanks the Spanish MEC for the BIO2006-15318 project. K.N.T, O.V.G., and P.N.G acknowledge the Federal Ministry for Science and Education (BMBF) for a grant in the framework of the BiotechGenoMik program, and K.N.T. thanks the Fonds der Chemischen Industrie for generous support. Authors are deeply indebted to A. Yanenko for sampling Kolguev Island coastal water and to the captain and crew of Research Vessel Urania for their assistance in deep-sea sampling in the Mediterranean Sea and to J. Manuel Franco for statistical analyses.

Supporting Online Material
www.sciencemag.org/cgi/content/full/326/5950/252/DC1
Materials and Methods
SOM Text
Figs. S1 to S14
Tables S1 to S6
References
Movie S1

26 March 2009; accepted 19 August 2009
10.1126/science.1174094

Unbiased Reconstruction of a Mammalian Transcriptional Network Mediating Pathogen Responses

Ido Amit,[1,2,3,4] Manuel Garber,[1]* Nicolas Chevrier,[2,3]* Ana Paula Leite,[1,5]* Yoni Donner,[1]* Thomas Eisenhaure,[2,3] Mitchell Guttman,[1,4] Jennifer K. Grenier,[1] Weibo Li,[2,3] Or Zuk,[1] Lisa A. Schubert,[6] Brian Birditt,[6] Tal Shay,[1] Alon Goren,[1,7] Xiaolan Zhang,[1] Zachary Smith,[1] Raquel Deering,[2,3] Rebecca C. McDonald,[2,3] Moran Cabili,[1] Bradley E. Bernstein,[1,3,7] John L. Rinn,[1] Alex Meissner,[1] David E. Root,[1] Nir Hacohen,[1,2,3]†‡ Aviv Regev[1,4,8]‡

Models of mammalian regulatory networks controlling gene expression have been inferred from genomic data but have largely not been validated. We present an unbiased strategy to systematically perturb candidate regulators and monitor cellular transcriptional responses. We applied this approach to derive regulatory networks that control the transcriptional response of mouse primary dendritic cells to pathogens. Our approach revealed the regulatory functions of 125 transcription factors, chromatin modifiers, and RNA binding proteins, which enabled the construction of a network model consisting of 24 core regulators and 76 fine-tuners that help to explain how pathogen-sensing pathways achieve specificity. This study establishes a broadly applicable, comprehensive, and unbiased approach to reveal the wiring and functions of a regulatory network controlling a major transcriptional response in primary mammalian cells.

Regulatory networks controlling gene expression serve as decision-making circuits within cells. For example, when immune dendritic cells (DCs) are exposed to viruses, bacteria, or fungi, they respond with transcriptional programs that are specific to each pathogen (*1*) and are essential for establishing appropriate immunological outcomes (*2*). These responses are initiated through specific receptors, such as Toll-like receptors (TLRs), that distinguish broad pathogen classes and are propagated through well-characterized signaling cascades (*2*). However, little is known about how the transcriptional network is wired to produce specific outputs.

Two major observational strategies have associated regulators with their putative targets on a genome scale (*3*): Cis-regulatory models rely on the presence of predicted transcription factor binding sites in the promoters of target genes (*3*–*5*), whereas trans-regulatory models are based on correlations between regulator and target expression (*3*–*6*). Because promoter binding sites and correlated expression are weak predictors of functional regulator-target linkages, such approaches are limited in their ability to produce reliable models of transcriptional networks (*3*). A complementary strategy is to systematically perturb every regulatory input and measure its effect on the expression of gene targets. This strategy has been successfully used in yeast (*7*–*9*) and sea urchin (*10*), but not in mammals.

A perturbation-based strategy for network reconstruction. We developed a perturbation strategy for reconstructing transcriptional networks in mammalian cells and used it to determine a network controlling the responses of DCs to pathogens (Fig. 1). First, we profiled gene expression at nine time points after stimulation with five pathogen-derived components and identified specific and shared genes that respond to each stimulus (fig. S1A). We used these profiles to identify 144 candidate regulators whose expression changed in response to at least one stimulus (*11*) (fig. S1B, top). We also identified a signature of 118 marker genes (fig. S1B, bottom) that captures the complexity of the response. We generated a validated lentiviral short hairpin RNA (shRNA) library for 125 of the 144 candidate regulators (fig. S1C, top), used it to systematically perturb each of the regulators in DCs, stimulated the cells with a pathogen component, and profiled the expression of the 118-gene signature (*12*) (fig. S1C, bottom). Finally, we used the measurements from the perturbed cells to derive a validated model of the regulatory network (fig. S1D).

Gene expression programs in response to TLR agonists. We measured genome-wide expression profiles in DCs exposed to PAM3CSK4 (PAM), a synthetic mimic of bacterial lipopeptides; polyinosine-polycytidylic acid [poly(I:C)], a viral-like double-stranded RNA; lipopolysaccharide (LPS), a purified component from Gram-negative *Escherichia coli*; gardiquimod, a small-molecule

[1]Broad Institute of MIT and Harvard, 7 Cambridge Center, Cambridge, MA 02142, USA. [2]Center for Immunology and Inflammatory Diseases, Massachusetts General Hospital, 149 13th Street, Charlestown, MA 02129, USA. [3]Harvard Medical School, Boston, MA 02115, USA. [4]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02142, USA. [5]Computational and Systems Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [6]NanoString Technologies, 530 Fairview Avenue N., Suite 2000, Seattle, WA 98109, USA. [7]Molecular Pathology Unit and Center for Cancer Research, Massachusetts General Hospital, Charlestown, MA 02129, USA. [8]Howard Hughes Medical Institute.
*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: nhacohen@partners.org
‡These authors contributed equally to this work.

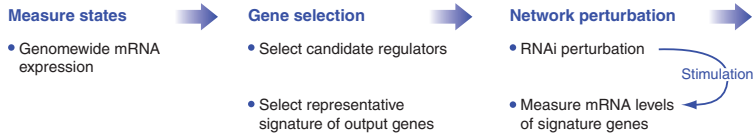

Fig. 1. A systematic strategy for network reconstruction. The strategy consists of four steps (left to right): state measurement using arrays; selection of regulators and response signatures; network perturbation with shRNAs against each regulator, followed by measurement of signature genes; and network reconstruction from the perturbational data.

agonist; and CpG, a synthetic single-stranded DNA. These compounds are known agonists of TLR2, TLR3, TLR4, TLR7, and TLR9, respectively. Poly(I:C) also activates the cytosolic viral RNA sensor MDA-5, and LPS can also act through co-receptors such as CD14; we therefore refer to the ligands rather than their receptors for clarity. On the basis of pilot experiments (fig. S2) (*11*), we measured mRNA expression at 0.5, 1, 2, 4, 6, 8, 12, 16, and 24 hours after stimulation with these pathogen components.

The observed transcriptional responses were classified into a "PAM-like" program and a "poly(I:C)-like" program, as well as a shared response [24.5% shared by PAM, poly(I:C), and LPS]. The LPS response (Fig. 2 and fig. S3) was largely the union of the "PAM-like" and "poly(I:C)-like" programs. This is partly explained by the known signaling pathways activated by these agonists. PAM binds TLR2 and signals through the MYD88 pathway; poly(I:C) binds TLR3 and MDA-5 and signals mostly through the TRIF and IPS-1 pathways, respectively; and LPS binds TLR4 and co-receptors and uses the MYD88 and TRIF pathways (*13*). It is also consistent with the known induction of an antiviral response by poly (I:C) and LPS (*14*). The "PAM-like" program is enriched for targets of the transcription factor NF-κB and for inflammatory responsive genes ($P < 6.1 \times 10^{-8}$), whereas the "poly(I:C)-like" program is enriched for interferon regulatory factors (IRFs) and for viral- and interferon-responsive genes (ISGs; $P < 8.3 \times 10^{-24}$). We thus term them the "inflammatory-like" and "antiviral-like" programs. A small number of genes are specific to a single stimulus. For example, ~250 genes are poly(I:C)-specific (1250 are shared with LPS), including several type I interferons (e.g., IFNA2, IFNA4; Fig. 2A). Surprisingly, 82% of the gardiquimod (TLR7) and CpG (TLR9) response was shared with the LPS response, but with a weaker antiviral component (fig. S4). This observation is unexpected given their different signaling mechanisms (*15*), but is highly reproducible and robust (fig. S4) (*11*).

Identification of candidate regulators and a response signature. To select potential regulators that mediate the observed transcriptional response, we focused on regulator genes whose expression changes during pathogen sensing [a reasonable assumption for many mammalian responses (*16*, *17*), including pathogen sensing (*1*, *4*)]. First, we reconstructed an observational trans-model of gene regulation (fig. S1B, top, and fig. S5A) (*11*) that associated 80 modules of co-regulated genes with 608 predictive regulators (fig. S5B) (*4*, *11*, *18*, *19*) automatically chosen out of a curated list of 3287 candidate regulators (*11*). Filtering identified 117 regulators above a minimal expression signal in at least one experiment (fig. S5B). These included known regulators from the NF-κB, STAT, and IRF families, as well as unexpected candidates such as the circadian regulator Timeless and the DNA methyltransferase Dnmt3a. Second, we added five constitutively expressed regulators whose cis-regulatory elements are enriched in the responsive genes (*11*). Third, to capture delayed responses or nonlinear relations, we incorporated 22 regulators whose expression changed by at least a factor of 2. This resulted in 144 candidate regulators, with a distribution of expression patterns similar to the general response (figs. S6 to S8 and table S1). The regulators' expression under LPS was conserved between DCs and functionally similar macrophages (Pearson correlation $r = \sim 0.9$ at 1 hour; fig. S9A) as well as between human macrophages and mouse DCs ($r = \sim 0.6$ at 2 hours; fig. S9B), supporting the functional relevance of the regulators' transcription.

To identify highly informative reporter genes for monitoring the effects of perturbing regulators, we devised GeneSelector (fig. S10A and table S2) (*11*). GeneSelector incrementally chooses genes (from our full expression data set) whose expression profile improves our discrimination of stimuli given the previously chosen genes. Using this approach, we identified the optimal time point (6 hours after activation; fig. S10B) and a set of 81 genes that distinguished the stimuli (*11*). We added 37 candidate regulators with

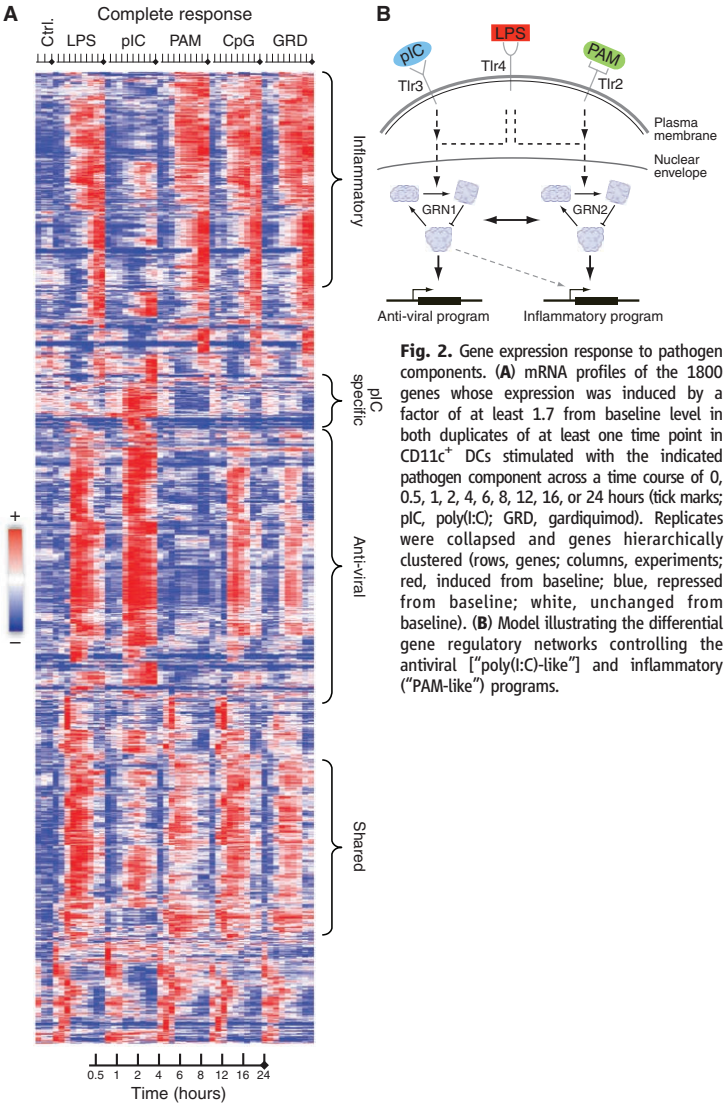

Fig. 2. Gene expression response to pathogen components. (**A**) mRNA profiles of the 1800 genes whose expression was induced by a factor of at least 1.7 from baseline level in both duplicates of at least one time point in $CD11c^+$ DCs stimulated with the indicated pathogen component across a time course of 0, 0.5, 1, 2, 4, 6, 8, 12, 16, or 24 hours (tick marks; pIC, poly(I:C); GRD, gardiquimod). Replicates were collapsed and genes hierarchically clustered (rows, genes; columns, experiments; red, induced from baseline; blue, repressed from baseline; white, unchanged from baseline). (**B**) Model illustrating the differential gene regulatory networks controlling the antiviral ["poly(I:C)-like"] and inflammatory ("PAM-like") programs.

detectable expression at the 6-hour time point, creating a signature of 118 genes. Finally, we added 10 control genes whose expression levels were unchanged under all stimuli, but whose (constant) basal levels varied from very low to high.

Perturbations, profiling, and modeling. We generated validated lentiviral shRNAs that knocked down the expression of 125 of our 144 candidate regulators in bone marrow DCs by at least 75% (fig. S11 and table S3) (*11*) and 32 shRNAs with no known gene targets as controls (figs. S12 and S13 and table S4) (*11*). To carry out our perturbational study, we selected a single treatment, LPS, that activates the majority of both the "inflammatory-like" and "antiviral-like" programs. After stimulation of shRNA-perturbed DCs with LPS for 6 hours, we used nCounter (*12*) to count transcripts of the 118 reporter and 10 control genes.

The changes in signature gene expression resulting from infection with each shRNA were used to construct a model that associated regulators to their targets. We expected increases in the transcript levels of reporter genes whose repressors are targeted by knockdown, and decreases in reporters whose activators are targeted. Our false discovery rate (FDR) model estimates the statistical significance of a change in transcripts in DCs infected with a given shRNA (*11*). We controlled for gene-specific noise by comparing to changes in the expression of each gene after perturbation with the control shRNAs (Fig. 3A), and for shRNA-specific noise by comparing to changes in the expression of the control genes after a given shRNA perturbation (Fig. 3B). We estimated the sensitivity of our calls from the 37 regulators, which are also included as target reporters (fig. S14) (*11*).

On the basis of these results, we identified a densely overlapping network with 2322 significant regulatory connections, including 1728 activations and 594 repressions (Fig. 3B, red and blue, respectively, at 95% confidence; tables S5 to S7). Of the 125 tested regulators, we confidently identified 100 with at least four targets. Among those were 24 hub regulators that were predicted to regulate more than 25% of the 118 genes measured, as well as 76 specific regulators each affecting the expression of 4 to 25 genes. On average, ~14 (±8; SD) regulators activated a target gene, and 5 (±5.8) regulators repressed it. Indirect effects may account for the large number of regulators we observed for each target.

Our perturbational model captured known regulatory features of the response, but also identified novel regulators. The reporter genes partitioned into two main clusters according to their response to perturbations (Fig. 3B and fig. S15A), consistent with the expression data: the "antiviral poly(I:C)–like" program reporters (e.g., Cxcl10, Isg15, and Ifit1) and the "inflammatory (PAM)–like" program reporters [e.g., IL1b, Cxcl2, IL6, and IL12b). We also found many known regulatory relations—for example, the NF-κB family of transcription factors (Rel, Rela, Relb, Nfkb1, Nfkb2, and Nfkbiz) regulating their known inflammatory gene targets. Our network provided evidence for the involvement of at least 68 additional regulators in the response to pathogens, of which 11 were hubs not previously associated with this system. Interestingly, 12 of the identified regulators (e.g., Hhex, Fus, Bat5, and Pa2g4) are in linkage disequilibrium with single-nucleotide polymorphisms (SNPs) associated with autoimmune and related diseases in genome-wide association studies (table S8).

The core inflammatory and antiviral programs. We next addressed how each regulator contributes to the generation of specific cell states. We first automatically defined the two major states induced by the five pathogen components with the use of non-negative matrix factorization (NMF) (*20*) and the original array data (*11*). This procedure identified two major expression components (termed "metagenes"): one predominantly determined by genes from the "inflammatory-like" program and the other by genes from the "antiviral-like" programs (Fig. 2A). Next, we quantified the effects of each regulator's knockdown on these two states (Fig. 3B, fig. S15A, and table S9) by classifying the nCounter expression measurements after a regulator's perturbation (*20*, *21*).

Finally, we used a regulator ranking score (*11*) to assign 33 (8 known) genes as regulators of the inflammatory state and 33 (15 known) genes as regulators of the antiviral state. This accurately classified the known activators of the inflammatory response (e.g., the NF-κB factors Rela, Nfkbiz, and Nfkb1; Fig. 3C, yellow in the inflammatory metagene) and of the antiviral response (e.g., Stat1, Stat2, Stat4, Irf8, and Irf9; Fig. 3C, yellow in the viral metagene). Although all perturbation experiments were conducted only under LPS stimulation (a bacterial component), we correctly classified factors known to mediate the response to other stimuli. Because 34 additional regulators were associated with both responses, it is possible that a single regulator can control genes in either state, depending on the differential timing of regulator activation, its level, or combinatorial regulation. Notably, for 12 of the transcription factors examined, we found an enriched cis-regulatory element in the appropriate metagene (*11*).

On the basis of the NMF scores (table S9), we identified an inflammatory subnetwork (fig. S15B), an antiviral subnetwork (Fig. 4A and fig. S15C), and several fine-tuning subnetworks that affect smaller numbers of genes from both responses (figs. S15D and S16) (*11*). The inflammatory subnetwork (fig. S15B) consisted of three regulatory modes: dominant activators (Cebpb, Bcl3, and Cited2) that induce more inflammatory targets than antiviral ones; cross-inhibitors (Nfkbiz, Nfkb1, Atf4, and Pnrc2) that induce inflammatory genes while repressing antiviral ones; and specific activators (Runx1 and Plagl2) that almost entirely target inflammatory genes. We observed that dominant activators mostly regulate effectors, whereas regulators are primarily controlled by cross-inhibitors.

Focusing on the network architecture, we found multiple feedforward circuits in this response, where an upstream regulator controls a target gene both directly and indirectly through a secondary regulator (*22*) (e.g., Fig. 4B and tables S10 and S11). The majority (76%, 4892 of 6444) of these feedforward circuits were found to be coherent (*22*), having the same direct and indirect effect on the regulated gene. The vast majority (80%) are type I loops (*23*) (Fig. 4B) with all-positive regulation (e.g., Nfkbiz activates E2f5 and both activate IL-6). Such feedforward circuits respond to persistent rather than transient stimulation, protecting the system from responding to spurious signals, as was shown for one circuit in LPS-stimulated macrophages (*24*). Our finding suggests that coherent feedforward loops, especially class I (*22*), are a general design principle in this system and may have a physiological impact on this response.

In the antiviral subnetwork, we identified a two-tiered regulatory circuit combining feedforward and feedback loops (Fig. 4A and table S11). This circuit has at the top the antiviral regulators Stat1 and Stat2, which regulate a full complement of antiviral reporters. The second-tier regulators Timeless, Rbl1, and Hhex are controlled by Stat1 and Stat2 and most likely form coherent feedforward loops that target specific subsets of genes. Timeless, Rbl1, and Hhex also feed back and promote the expression of the Stat regulators. This circuit is repressed through the cell cycle regulator and RNA binding protein Fus (*25*), acting as a single dominant inhibitor of 43 viral genes.

Finally, we derived a core network incorporating the regulators with the most substantial impact on each response, on the basis of the number, magnitude, and logic of targets that each regulator affects (*11*). The core network (Fig. 4C) has 24 regulators, 13 of which have previously been identified as key factors regulating the inflammatory or antiviral responses; the other 11 have not been previously implicated in either response. Of these 24 regulators, 19 are transcription factors, three are chromatin modifiers, and two are RNA binding proteins. The regulators apparently distinguish the two programs through cross-inhibition (Fig. 4C, gray lines) or dominant activation (Fig. 4C). The core network also explains how differential expression of secreted factors is specified, leading to the activation and migration of appropriate cell types for different pathogens (*11*, *26*) (fig. S17).

Embedded within the many known regulators of the antiviral response (Fig. 4C and fig. S15C), we found a large set of regulators not previously associated with this response. These included several known regulators of the cell cycle and the circadian rhythm, including Rbl1, Jun, Rb, E2f5, E2f8, Nmi, Fus, and Timeless, several of which were placed in our core network. This

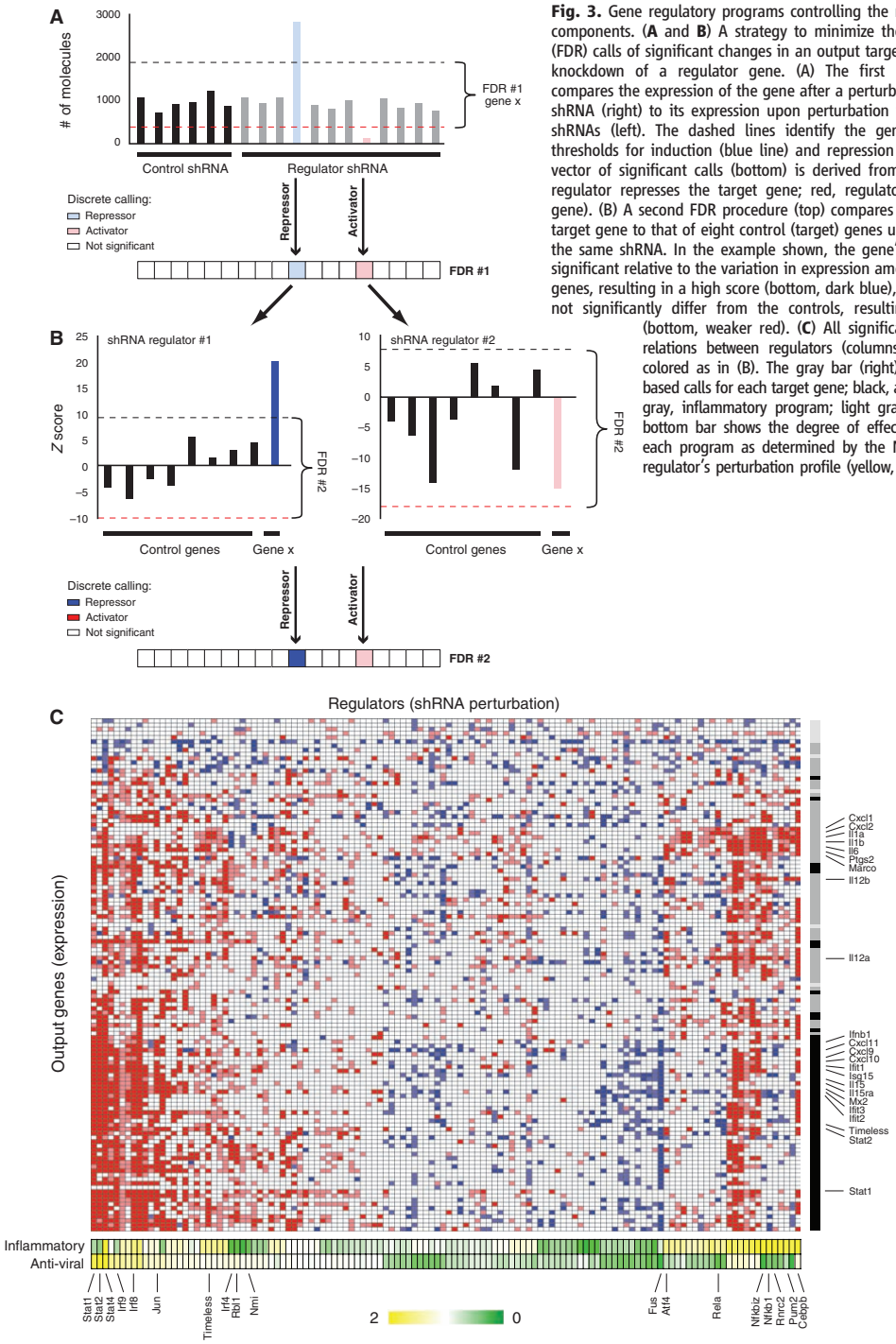

Fig. 3. Gene regulatory programs controlling the response to pathogen components. (**A** and **B**) A strategy to minimize the false discovery rate (FDR) calls of significant changes in an output target gene resulting from knockdown of a regulator gene. (A) The first FDR procedure (top) compares the expression of the gene after a perturbation with a regulator shRNA (right) to its expression upon perturbation with 32 nontargeting shRNAs (left). The dashed lines identify the gene-specific FDR-based thresholds for induction (blue line) and repression (red line). A discrete vector of significant calls (bottom) is derived from the raw data (blue, regulator represses the target gene; red, regulator induces the target gene). (B) A second FDR procedure (top) compares the expression of the target gene to that of eight control (target) genes upon perturbation with the same shRNA. In the example shown, the gene's induction (left) was significant relative to the variation in expression among the control target genes, resulting in a high score (bottom, dark blue), but its repression did not significantly differ from the controls, resulting in a lower score (bottom, weaker red). (**C**) All significant (95% confidence) relations between regulators (columns) and targets (rows), colored as in (B). The gray bar (right) represents the NMF-based calls for each target gene; black, antiviral program; dark gray, inflammatory program; light gray, control genes. The bottom bar shows the degree of effect by the regulator on each program as determined by the NMF projection of the regulator's perturbation profile (yellow, high; green, low).

suggests that a cell cycle regulatory circuit was co-opted to function in the antiviral response in DCs (with no observable effect on cell cycle progression; fig. S18). Because we identified these antiviral regulatory relations in perturbation experiments using DCs stimulated with the bacterial component LPS, we silenced four regulators (Timeless, Rbl1, Jun, and Nmi) after exposure to the viral component poly(I:C). Each of the four regulators had a strong impact on the antiviral program, more than was observed under LPS stimulation (Fig. 4D), and on affected genes (e.g., type I IFNs) whose expression is poly(I:C)-specific. Nmi affected a smaller set of genes, consistent with the model's prediction. These results demonstrate our ability to correctly predict function in unobserved conditions.

Although most antiviral genes are induced after stimulation with the bacterial component LPS, a few critical ones are expressed specifically in poly(I:C) stimulation or follow distinct patterns in each stimulus. For example, in response to viral infection, cells induce the production of interferon β1 (Ifnb1), a crucial mediator of the antiviral response. Because high levels of Ifnb1 may be deleterious to the host if infected by specific bacteria (*27*), we predicted that specific mechanisms insulate Ifnb1's regulation from the response to LPS. Indeed, although Ifnb1 expression was induced in the first 2 hours of stimulation with LPS, this expression declined at subsequent time points, in contrast to its sustained induction after poly(I:C) treatment (Fig. 5A). Our model suggested that three regulators known to affect chromatin remodeling (*25*, *28*, *29*) are Ifnb1 repressors in LPS (Fig. 5B): the Polycomb complex subunit Cbx4 (*28*), Fus (*25*), and the DNA methyltransferase Dnmt3a (*29*). Cbx4 appeared to confer antiviral specificity to Ifnb1 induction because it was induced within the first 2 hours of PAM and LPS treatment but not by poly(I:C) (Fig. 5C), and Cbx4 knockdown caused induction of Ifnb1 mRNA and protein during LPS treatment (Fig. 5D and fig. S19A) but had no effect on the induction of the chemokine Cxcl10, a poly(I:C)- and LPS-induced gene (fig. S19B). Cbx4 knockdown did not affect Ifnb1 during PAM activation (Fig. 5E), when the antiviral response is not induced. Combined with evidence for chromatin changes around the Ifnb1 locus and its closest neighbor gene, Ptplad2 (fig. S20A), which has a similar dependence on Cbx4, these data are consistent with an effect by Cbx4 on local chromatin organization (fig. S20, B and C). Cbx4 knockdown affected few genes (~120 up-regulated and ~120 down-regulated genome-wide; table S12). Because most up-regulated

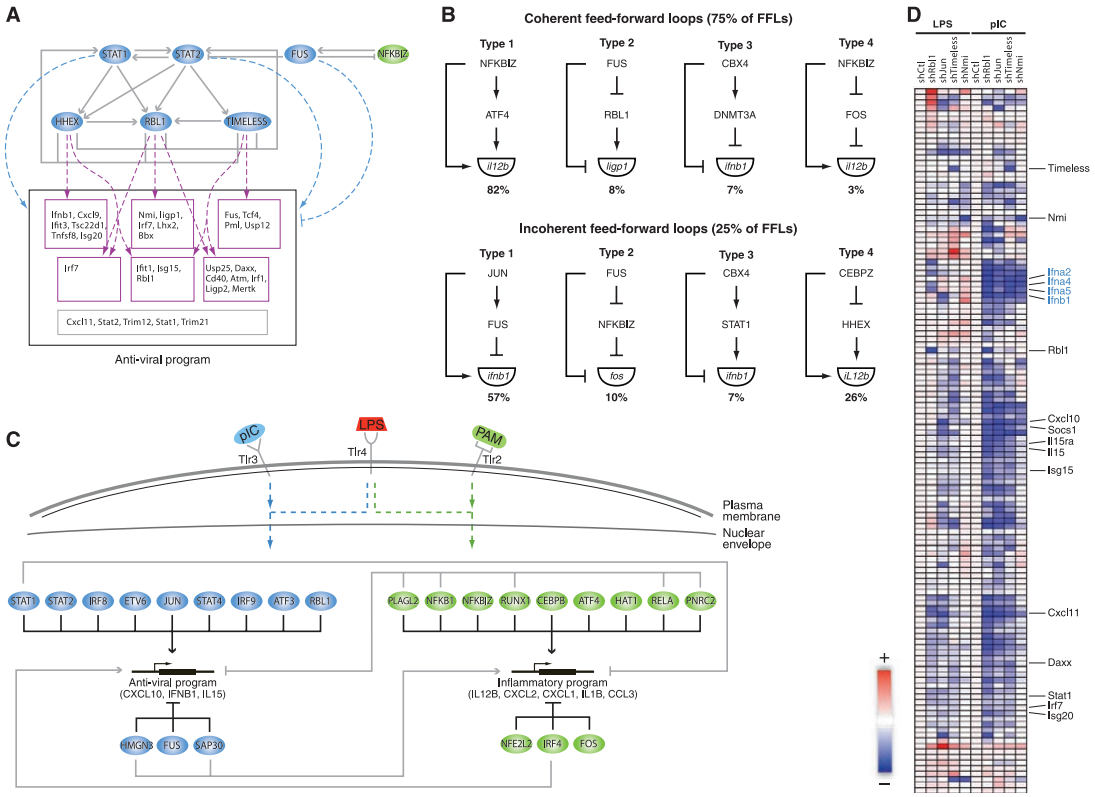

Fig. 4. The core regulatory circuits controlling the inflammatory and antiviral responses. **(A)** The antiviral subnetwork shows regulatory relations between the core antiviral regulators (blue nodes, top), their targets (boxes, bottom), each other, and inflammatory regulators (green node, top right). The two top regulators, Stat1 and Stat2, activate all antiviral targets (dashed blue arrows). The second-tier regulators activate subsets of targets (dashed purple arrows). **(B)** Examples of feedforward loop classes identified in the network, with fraction of each class. **(C)** A core regulatory network of the inflammatory and antiviral programs, consisting of the most distinct regulators, and their relation to ligands and receptors (top). Pointed arrows, induction; blunt arrows, repression; green ovals, inflammatory regulators; blue ovals, antiviral regulators. Example target genes are noted. **(D)** nCounter expression profiles for the target genes (rows) upon perturbation with shRNAs against a subset of viral regulators (columns) and followed by stimulation with LPS (left) or poly(I:C) (right). All values are normalized by expression in cells infected with a control shRNA and under the same stimulus (shCtl).

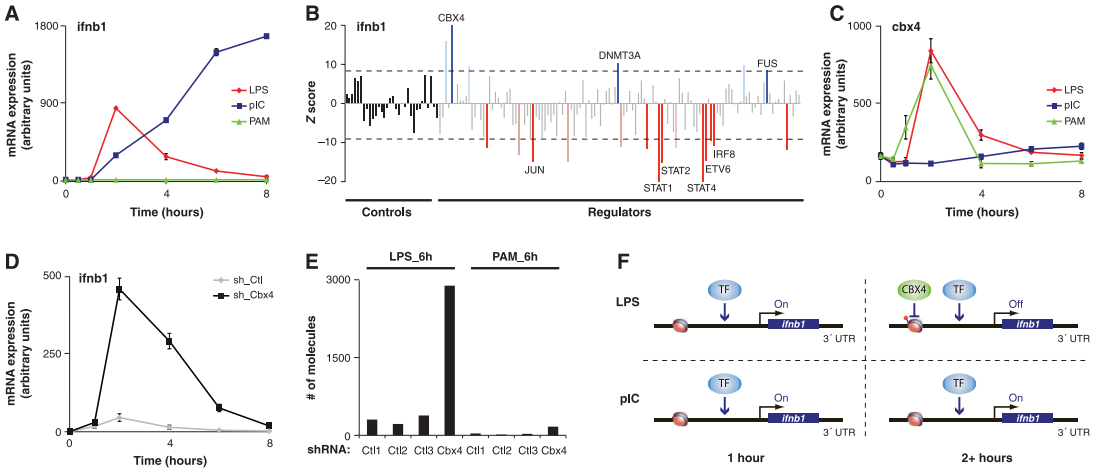

Fig. 5. The polycomb component Cbx4 selectively restricts Ifnb1 production under bacterial perturbations. (**A** and **C**) LPS (red), poly(I:C) (blue), and PAM (green)–induced expression of Ifnb1 (A) and Cbx4 (C) derived from data in Fig. 2A. (**B**) Ifnb1 expression (by nCounter) in response to LPS in DCs perturbed by control shRNAs or shRNAs targeting each of 125 regulators (format as in Fig. 3B). (**D**) Ifnb1 mRNA levels (by quantitative reverse transcription polymerase chain reaction) after LPS treatment in unsorted mouse bone marrow DCs perturbed with an shRNA against Cbx4 (black) or a control shRNA (gray); signals are relative to $t = 0$. (**E**) Ifnb1 mRNA levels (by nCounter) at 6 hours after exposure to LPS or PAM in bone marrow DCs perturbed with an shRNA against Cbx4 or one of three control shRNAs. (**F**) Model for bacterial-specific repression of Ifnb1 by Cbx4. Both poly(I:C) and LPS induce Ifnb1 expression early (left), but only LPS induces Cbx4, which then represses the *Ifnb1* locus at a later time (right, top).

genes show a precise temporal pattern in unperturbed cells akin to that of Cbx4—they are induced quickly and return to basal level by 2 to 4 hours (fig. S21, A to F)—we conclude that a chromatin modifier can act like a transcription factor controlling the precise expression of specific genes in the regulatory program.

Taken together, our results suggest a model of a transcriptional negative feedback loop, controlling Ifnb1 expression in LPS stimulation, wherein the induced proinflammatory regulator and chromatin modifier Cbx4 represses transcription by modifying the chromatin in the Ifnb1 locus, generating the specificity needed to drive the inflammatory versus the antiviral response (Fig. 5F). The type III coherent feedforward loop formed by Cbx4 and Dnmt3a (Fig. 4B) is consistent with a delayed repression of Ifnb1. Because neither regulator carries a sequence-specific DNA binding domain, the factors responsible for their guidance to the *Ifnb1* locus remain unknown.

Discussion. A central goal of our study was to address the mechanistic basis for pathogen-specific responses. Consistent with previous studies (*14*), we distinguished two key programs, a PAM (TLR2)–like inflammatory response and a poly(I:C) (TLR3/MDA-5)–like antiviral response, which are together induced by LPS, a Gram-negative bacterial component and a TLR4 ligand. These programs reflect both qualitative and quantitative differences between the required functional responses, and are consistent with the cross-protection between certain bacteria and virus infections (*14*). The broad effect of LPS allowed us to focus on a single stimulus and time point, but screens with other stimuli may identify additional unique regulators.

We found that these two responses are controlled by two corresponding regulatory arms, uncovering a mechanistic basis for the observed transcriptional responses. These two arms are integrated into a core network of 24 regulators that balance specific and shared responses through dominant activation and cross-inhibition. In the inflammatory response, we found several feed-forward loops, which may ensure response to only persistent and not sporadic signals. In the antiviral response, we discovered a two-tiered circuit involving feedback and feedforward loops, implicating a module of cell cycle regulators (Jun, Rbl1, Timeless, and Nmi), which we directly validated. More than 75 additional genes work to further fine-tune the regulation of gene targets. This perturbational model identifies many regulatory relations that would have been missed by nonsystematic approaches.

Our work establishes an unbiased, straightforward, and general framework for network reconstruction in mammalian cells (*11*), including several strategies to leverage shRNA for the study of gene regulation. This approach can be executed at substantial scale and reasonable cost, and is compatible with the challenge of deciphering the multiple regulatory systems that operate in mammals. It can be expanded to derive increasingly detailed models and to distinguish direct from indirect targets.

Our study will facilitate the development of new computational approaches to infer regulatory models. Although many computational approaches have attempted to derive observational models, their quality has been difficult to evaluate (*3*). The data generated here include expression profiles for training a model, as well as a perturbational unbiased screen for testing its quality (ftp://ftp.broadinstitute.org/pub/papers/dc_network). When we compared the perturbational model to our observational model, we found that many candidate regulators were correctly identified in both (figs. S5 and S22). However, there were also numerous false positive relations in the observational model, attributable to the fact that both the correct regulator and many others have indistinguishable expression (figs. S22 and S23).

The high-resolution map we constructed has important biomedical implications. By identifying regulators that mediate the differential control of specific gene pairs (e.g., IL-23 versus IL-12, fig. S17) and entire regulatory arms (e.g., viral versus inflammatory), it opens the way for therapeutic targeting of specific pathways to control disease or enhance vaccine efficacy. Furthermore, 12 of our regulators reside in genetic loci that were in linkage disequilibrium with SNPs associated with autoimmune and related diseases. The identified genes and their impact on DCs provide hypotheses to help explain how alleles of genes in a cascade may alter susceptibility to specific infections or immune disorders in humans.

References and Notes

1. Q. Huang *et al.*, *Science* **294**, 870 (2001).
2. T. Kawai, S. Akira, *Int. Immunol.* **21**, 317 (2009).

3. H. D. Kim, T. Shay, E. K. O'Shea, A. Regev, *Science* **325**, 429 (2009).
4. S. A. Ramsey *et al.*, *PLOS Comput. Biol.* **4**, e1000021 (2008).
5. H. Suzuki *et al.*, *Nat. Genet.* **41**, 553 (2009).
6. E. Segal *et al.*, *Nat. Genet.* **34**, 166 (2003).
7. A. P. Capaldi *et al.*, *Nat. Genet.* **40**, 1300 (2008).
8. Z. Hu, P. J. Killion, V. R. Iyer, *Nat. Genet.* **39**, 683 (2007).
9. C. T. Workman *et al.*, *Science* **312**, 1054 (2006).
10. D. H. Erwin, E. H. Davidson, *Nat. Rev. Genet.* **10**, 141 (2009).
11. See supporting material on *Science* Online.
12. G. K. Geiss *et al.*, *Nat. Biotechnol.* **26**, 317 (2008).
13. J. C. Kagan *et al.*, *Nat. Immunol.* **9**, 361 (2008).
14. S. Doyle *et al.*, *Immunity* **17**, 251 (2002).
15. T. Kawai *et al.*, *Nat. Immunol.* **5**, 1061 (2004).
16. I. Amit *et al.*, *Nat. Genet.* **39**, 503 (2007).
17. D. Pe'er, A. Regev, A. Tanay, *Bioinformatics* **18** (suppl. 1), S258 (2002).
18. S. I. Lee *et al.*, *PLoS Genet.* **5**, e1000358 (2009).
19. H. Zou, T. Hastie, *J. R. Stat. Soc. B* **67**, 301 (2005).
20. J. P. Brunet, P. Tamayo, T. R. Golub, J. P. Mesirov, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 4164 (2004).
21. J. P. Daily *et al.*, *Nature* **450**, 1091 (2007).
22. S. Mangan, U. Alon, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 11980 (2003).
23. U. Alon, *Nat. Rev. Genet.* **8**, 450 (2007).
24. V. Litvak *et al.*, *Nat. Immunol.* **10**, 437 (2009).
25. X. Wang *et al.*, *Nature* **454**, 126 (2008).
26. A. D. Luster, *Curr. Opin. Immunol.* **14**, 129 (2002).
27. T. Decker, M. Muller, S. Stockinger, *Nat. Rev. Immunol.* **5**, 675 (2005).
28. E. Bernstein *et al.*, *Mol. Cell. Biol.* **26**, 2560 (2006).
29. B. Li *et al.*, *Biochem. J.* **405**, 369 (2007).
30. We thank E. Lander, I. Wapinski, D. Pe'er, N. Friedman, J. Kagan, A. Luster, V. Kuchroo, and A. Citri for discussions and comments; L. Gaffney for assistance with artwork; S. Gupta and the Broad Genetic Analysis Platform for microarray processing; and T. Mikkelsen and the Broad Sequencing Platform for help with the ChIP-seq experiments. Supported by the Human Frontier Science Program Organization and a Claire and Emanuel G. Rosenblatt Award from the American Physicians Fellowship for Medicine in Israel (I.A.); NIH grant R21 AI71060 and the NIH New Innovator Award (N.H.); and a Career Award at the Scientific Interface from the Burroughs Wellcome Fund, an NIH Pioneer Award, and the Sloan Foundation (A.R.). A.R. is an Early Career Scientist of the Howard Hughes Medical Institute and an investigator of the Merkin Foundation for Stem Cell Research at the Broad Institute. Complete microarray data sets are available at the Gene Expression Omnibus (accession no. GSE17721).

Supporting Online Material

www.sciencemag.org/cgi/content/full/1179050/DC1
Materials and Methods
SOM Text
Figs. S1 to S23
Tables S1 to S16
References

14 July 2009; accepted 26 August 2009
Published online 3 September 2009;
10.1126/science.1179050
Include this information when citing this paper.

Mapping Excited-State Dynamics by Coherent Control of a Dendrimer's Photoemission Efficiency

Daniel G. Kuroda,[1]* C. P. Singh,[1]† Zhonghua Peng,[2] Valeria D. Kleiman[1]‡

Adaptive laser pulse shaping has enabled impressive control over photophysical processes in complex molecules. However, the optimal pulse shape that emerges rarely offers straightforward insight into the excited-state properties being manipulated. We have shown that the emission quantum yield of a donor-acceptor macromolecule (a phenylene ethynylene dendrimer tethered to perylene) can be enhanced by 15% through iterative phase modulation of the excitation pulse. Furthermore, by analyzing the pulse optimization process and optimal pulse features, we successfully isolated the dominant elements underlying the control mechanism. We demonstrated that a step function in the spectral phase directs the postexcitation dynamics of the donor moiety, thus characterizing the coherent nature of the donor excited state. An accompanying pump-probe experiment implicates a 2+1 photon control pathway, in which the optimal pulse promotes a delayed excitation to a second excited state through favorable quantum interference.

Since ancient times, humans have been trying to control the transformation of matter. For more than a century, absorption of light has been used to initiate photochemical reactions. It is only in the past 20 years, however, that researchers have devised techniques to steer the ensuing dynamics through modulation of the optical excitation field. Such quantum, or coherent, control schemes (*1*–*3*) use laser-derived electric fields to direct the motion of wave packets along excited-state potential energy surfaces (*4*–*6*). In principle, the phases and amplitudes in the applied field necessary to achieve a given outcome can be obtained from the field's coupling to the molecular Hamiltonian. In practice, rational design of the requisite pulse shapes remains an insurmountable problem for large molecules in condensed phase: The complete Hamiltonian is either unknown or too complex to be used in electric field calculations. Instead, researchers have relied on empirical methods whereby pulse shapes are determined through iterative optimization using the desired product (e.g., fluorescence quantum yield) as a feedback parameter (*7*). Thus, photoinduced processes can be actively manipulated without previous knowledge of the Hamiltonian and the light-matter couplings (*8*–*11*).

This closed-loop feedback technique has proven powerfully versatile. For example, modulation of isomerization yield in the natural photoreceptor bacteriorhodopsin (*11*), and control of quantum efficiency in both natural and artificial photosynthetic antennae by tuning spectral amplitude and phase in resonant linear excitation (*9*, *12*), were shown. Nonetheless, very few closed-loop experiments have yielded optimal pulse shapes that can be directly explained in terms of known molecular properties of the system under investigation (*9*, *11*, *13*–*15*). Major hindrances to attain such insight are the often intricate relationship between the variables to be optimized and the molecular response; the large number of parameters used for the generation of arbitrarily modulated pulses; and the often arbitrary, random pathways to optimization generated by the iteration algorithms. A major goal in the field is thus to develop a procedure for gleaning molecular insight from coherent control, especially in systems where the photophysical or photochemical pathways are presently unknown. For example, using optimal control, Branderhorst *et al.* (*15*) were able to identify wave packets with minimum position variance as candidates to minimize coupling to the bath and thus increase coherence robustness.

Our approach toward this end is to search for a confined set of parameters that directly govern the optimization. The idea is to express all the independent electric field parameters obtained from the collection of closed-loop optimization data as a combined set, filtering out those variables with redundant or negligible effects on the molecular response (*16*–*20*). Ideally, it ought to be more straightforward to associate these fewer parameters with a physical property.

We apply this approach to coherent control studies of the emissive properties of the phenylene ethynylene dendrimer $2G_2$-m-Per (Fig. 1) (*21*), designed to mimic natural light harvesting systems (*22*). Phenylene ethynylene dendrimers are rigid macromolecules with a high quantum yield for energy transfer from donor to ac-

[1]Department of Chemistry, Center for Chemical Physics, University of Florida, Gainesville, FL 32611–7200, USA. [2]Department of Chemistry, University of Missouri–Kansas City, Kansas City, MO 64110–2499, USA.

*Present address: Department of Chemistry, University of Pennsylvania, Philadelphia, PA 19104–6323, USA.
†On leave from Raja Ramanna. Centre for Advanced Technology, Indore, India.
‡To whom correspondence should be addressed. E-mail: kleiman@chem.ufl.edu

Cell

Systematic Discovery of TLR Signaling Components Delineates Viral-Sensing Circuits

Nicolas Chevrier,[1,2,3,4] Philipp Mertins,[1] Maxim N. Artyomov,[1,6] Alex K. Shalek,[7] Matteo Iannacone,[5,9] Mark F. Ciaccio,[8] Irit Gat-Viks,[1,6,10] Elena Tonti,[5,9] Marciela M. DeGrace,[1,2,3] Karl R. Clauser,[1] Manuel Garber,[1] Thomas M. Eisenhaure,[1,2] Nir Yosef,[1,6] Jacob Robinson,[7] Amy Sutton,[7] Mette S. Andersen,[7] David E. Root,[1] Ulrich von Andrian,[5] Richard B. Jones,[8] Hongkun Park,[7] Steven A. Carr,[1] Aviv Regev,[1,6,12] Ido Amit,[1,2,3,6,11,12,*] and Nir Hacohen[1,2,3,12,*]

[1]Broad Institute of MIT and Harvard, 7 Cambridge Center, Cambridge, MA 02142, USA
[2]Center for Immunology and Inflammatory Diseases, Massachusetts General Hospital, Charlestown, MA 02129, USA
[3]Department of Medicine
[4]Graduate Program in Immunology, Division of Medical Sciences
[5]Immune Disease Institute and Department of Pathology
Harvard Medical School, Boston, MA 02115, USA
[6]Howard Hughes Medical Institute, Department of Biology, MIT, Cambridge, MA 02142, USA
[7]Departments of Chemistry, Chemical Biology, and Physics, Harvard University, Cambridge, MA 02138, USA
[8]The Ben May Department for Cancer Research and the Institute for Genomics and Systems Biology, The University of Chicago, Chicago, IL 60637, USA
[9]Division of Immunology, Infectious Diseases and Transplantation, San Raffaele Scientific Institute, Milan 20132, Italy
[10]Present address: Department of Cell Research and Immunology, George S. Wise Faculty of Life Sciences, Tel Aviv University, Tel Aviv 69978, Israel
[11]Present address: Department of Immunology, Weizmann Institute of Science, Rehovot 76100, Israel
[12]These authors contributed equally to this work
*Correspondence: ido.amit@weizmann.ac.il (I.A.), nhacohen@partners.org (N.H.)
DOI 10.1016/j.cell.2011.10.022

SUMMARY

Deciphering the signaling networks that underlie normal and disease processes remains a major challenge. Here, we report the discovery of signaling components involved in the Toll-like receptor (TLR) response of immune dendritic cells (DCs), including a previously unkown pathway shared across mammalian antiviral responses. By combining transcriptional profiling, genetic and small-molecule perturbations, and phosphoproteomics, we uncover 35 signaling regulators, including 16 known regulators, involved in TLR signaling. In particular, we find that Polo-like kinases (*Plk*) 2 and 4 are essential components of antiviral pathways in vitro and in vivo and activate a signaling branch involving a dozen proteins, among which is *Tnfaip2*, a gene associated with autoimmune diseases but whose role was unknown. Our study illustrates the power of combining systematic measurements and perturbations to elucidate complex signaling circuits and discover potential therapeutic targets.

INTRODUCTION

Signaling networks detect and respond to environmental changes, and defects in their wiring can contribute to diseases. For example, Toll-like receptors (TLRs) sense microbial molecules and trigger signaling pathways critical for host defense (Takeuchi and Akira, 2010). Genetic defects in components of the TLR and other pathogen-sensing pathways have been linked to human diseases. Hence, rational targeting of these pathways should help in better manipulating immune responses associated with infections, autoimmunity, and vaccines (Hennessy et al., 2010).

However, despite extensive studies, many components of TLR and other biological networks are unknown, and many genes associated with disease have not been assigned to a function or a pathway. A key challenge is thus to systemically dissect mammalian signaling networks, by determining the functions of their components and placing them within pathways. Previously, we introduced an integrated experimental and computational approach to decipher the TLR transcriptional network of immune dendritic cells (DCs) (Amit et al., 2009), allowing us to identify transcriptional regulators and to define their impact on TLR responses in DCs. For example, we found a host of cell-cyle regulators—*Rbl1*, *Rb*, *Myc*, *Jun*, and *E2fs*—that are required for antiviral transcriptional responses in nondividing DCs.

Here, we adapt and expand this approach to the discovery and validation of TLR signaling components in DCs (Figure S1 available online). First, to identify candidate components, we rely on transcriptional feedbacks, whereby a signaling circuit regulates the transcript levels of genes encoding some, but not all, of its components (Amit et al., 2007; Fraser and Germain, 2009; Freeman, 2000). Second, we perturb these candidates with shRNAs and measure the effects on a representative

signature of >100 TLR-activated genes. Third, we use functional phosphoproteomics to expand the pathway's scope to components whose mRNA levels may be unchanged upon TLR activation. Applying this approach iteratively, we discovered 19 functional components, including a signaling arm mediated by two Polo-like kinases (*Plk2* and *4*) that participate in regulating well-established host antiviral pathways.

RESULTS

Transcripts for Signaling Components Are Regulated upon TLR Stimulation

To identify candidate components of pathogen-sensing pathways, we used genome-wide mRNA profiles, previously measured at 10 time points along 24 hr following stimulation of primary bone marrow-derived DCs (BMDCs) with lipopolysaccharide (LPS; TLR4 agonist), polyinosinic:polycytidylic acid (poly(I:C); recognized by TLR3 and the cytosolic viral sensor MDA-5), or Pam3CSK4 (PAM; TLR2 agonist) (Amit et al., 2009). These three TLRs activate transcriptional programs referred to here as "inflammatory" (TLR2), "antiviral" (TLR3), or both (TLR4) (Figure 1A) (Amit et al., 2009; Doyle et al., 2002).

Our analysis uncovered 280 genes annotated as known or putative signaling molecules that were differentially expressed following stimulation: 115 kinases, 69 phosphatases, and 96 other regulators, such as adaptors and scaffolds (Figure 1B, Table S1, and Experimental Procedures). These 280 genes were enriched for canonical pathways of the TLR network such as MAP kinase ($p < 1.22 \times 10^{-15}$, overlap 25/87, hypergeometric test), TLR (e.g., *Myd88*, *Traf6*, *Irak4*, *Tbk1*; $p < 8.43 \times 10^{-12}$, 21/86), and PI3K ($p < 2.58 \times 10^{-8}$, 11/33) pathways, as well as the PYK2 pathway ($p < 3.12 \times 10^{-10}$, 12/29), which was recently associated with the TLR system (Wang et al., 2010). Overall, 94 of the 280 genes (33%) were associated with the TLR network in the literature (Table S1), supporting the validity of our candidate selection strategy. The remaining 186 genes (67%) represent candidate TLR components. To test their putative function in TLR signaling, we selected a subset of 23 candidates based on their strong differential expression and to proportionally represent the five main induced expression clusters (Figures 1B and 1C). We also selected six canonical TLR components (*Myd88*, *Mapk9*, *Tbk1*, *Ikbke*, *Tank*, and *Map3k7*) as benchmarks (Figures 1A and 1D).

A Perturbation Strategy Places Uncharacterized Signaling Components within the Antiviral and Inflammatory Pathways

We perturbed our 6 positive controls and 17 of the 23 candidates in BMDCs using shRNA-encoding lentiviruses (6 candidates showed poor knockdown efficiency) (Table S1). We stimulated the cells with LPS and used a multiplex mRNA counting method to measure the effect of gene silencing on the mRNA levels of 118 TLR response signature genes, representing the inflammatory and antiviral programs (Figure 2A). Notably, the expression of the 118 genes was not affected in BMDCs transduced with lentivirus compared to untransduced cells (Amit et al., 2009). We determined statistically significant changes in the expression of signature transcripts upon individual knockdowns based on comparison to 10 control genes, whose expression remains unchanged upon TLR activation, and to control shRNAs (Experimental Procedures). Finally, we associated signaling molecules and downstream transcriptional regulators that may act in the same pathway by comparing the perturbational profiles of the 23 signaling molecules (6 canonical and 17 candidates) to each other and to those of the 123 transcriptional regulators (including transcription and chromatin factors and RNA-binding proteins) previously tested (Figures 2 and S2 and Table S2) (Amit et al., 2009).

Perturbing 5 of the 6 canonical signaling molecules strongly affected the expression of TLR signature genes, consistent with their known roles (Figure 2A and Table S2) and validating our approach. For example, perturbing *Myd88*, a known inflammatory adaptor, specifically abrogated the transcription of inflammatory genes (e.g., *Cxcl1*, *Il1a*, *Il1b*, *Ptgs2*, *Tnf*; Figure 2A), similar to perturbations of downstream inflammatory transcription factors (e.g., *Nfkb1*, *Nfkbiz*; Figure 2B). In addition, *Tank* acted as a negative regulator of a subset of antiviral genes (Figure 2A), as expected (Kawagoe et al., 2009), and *Tbk1* knockdown affected both antiviral and inflammatory outputs (Figure 2A), consistent with findings that *Tbk1* regulates NF-κB complexes (Barbie et al., 2009; Chien et al., 2006). Notably, *Ikbke* (IKK-ε) knockdown did not affect our gene signature, consistent with previous observations that IKK-$\varepsilon^{-/-}$ DCs respond normally to LPS and viral challenges (Matsui et al., 2006). Thus, IKK-ε may be either not functional or redundant in our system.

All of the 17 candidate signaling molecules tested, except *Plk2* (discussed below), affected at least 6 of the 118 genes (on average, 16.6 targets ± 10.4 standard deviation [SD]), and 12 affected more than 10% of the genes (Figures S2A and S2D). Notably, perturbations of these 17 candidates did not affect BMDC differentiation (88.3% ± 6.8% SD of $CD11c^+$ cells; Table S1). These effects are comparable to those of known signaling molecules and transcriptional regulators in this system (Figures S2B–S2E). For example, the receptor tyrosine kinase *Met*, not previously associated with TLR signaling, affected a number of signature genes similar to *Tbk1* (Figures S2C and S2D), in both the inflammatory and antiviral programs (Figure 2A). Conversely, both the phosphatase *Ptpre* and the adaptor *Socs6* positively regulated the inflammatory program, although negatively regulating some antiviral genes (Figure 2B). Of the 17 candidates tested when we originally conducted this screen, 10 have subsequently been reported in other studies as functional in the TLR system (Table S1), providing an independent confirmation. For example, here *Map3k8* knockdown affected both inflammatory and antiviral target genes (Figure 2A), consistent with its reported role in the TLR pathways based on *Sluggish* mice (Xiao et al., 2009).

We identified both primary (e.g., *Myd88*) and secondary (e.g., *Stat1*) mediators of TLR responses. Although secondary mediators are not part of the initial intracellular signaling cascade, they are important physiological components of the TLR response, and their pertubation can lead to phenotypic outcomes similar to those of primary components. For example, the receptor tyrosine kinase *Mertk* acted as both a positive and negative regulator of some inflammatory and antiviral genes (e.g., *Ifnb1*), respectively (Figure 2A), consistent with its reported role as a secondary inhibitor of the TLR pathways (Rothlin et al., 2007).

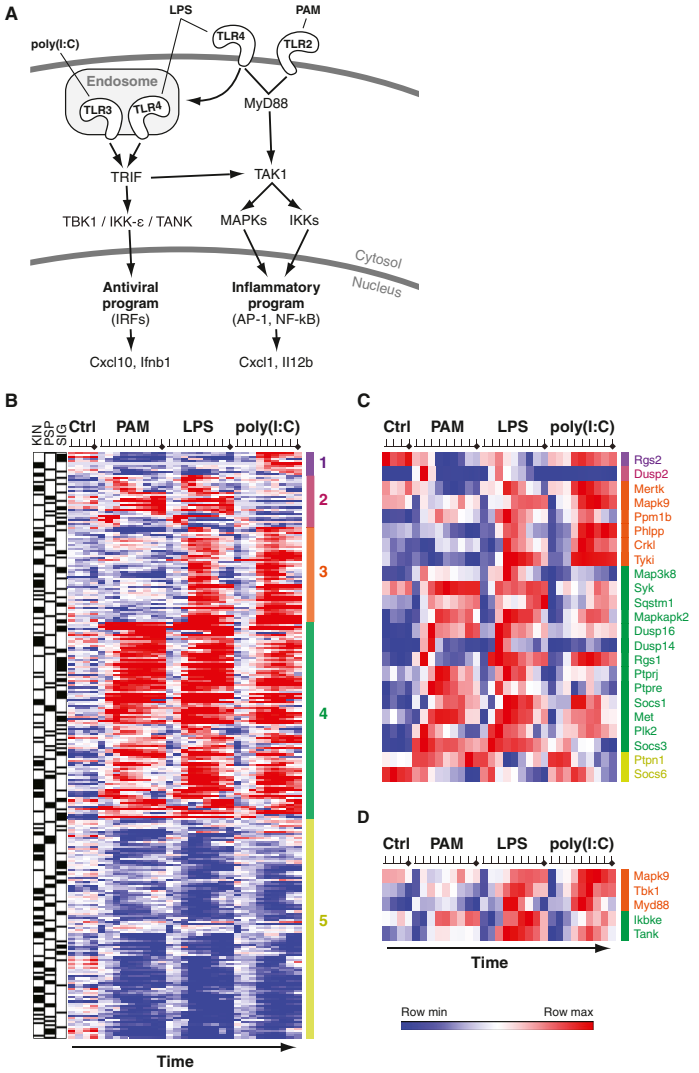

Figure 1. mRNAs of Signaling Components Are Differentially Regulated upon TLR Stimulation

(A) Simplified schematic of the TLR2, 3, and 4 pathways (Takeuchi and Akira, 2010).

(B) mRNA expression profiles of differentially expressed signaling genes. Shown are expression profiles for 280 differentially expressed signaling genes (rows) at different time points (columns): a control time course (no stimulation, Ctrl) and following stimulations with Pam3CSK4 (PAM), LPS, and poly(I:C). Tick marks: time point poststimulation (0.5, 1, 2, 4, 6, 8, 12, 16, 24 hr). Shown are genes with at least a 1.7-fold change in expression compared to prestimulation levels in both duplicates of at least one time point. The three leftmost columns indicate kinase (KIN), phosphatase (PSP), and signaling regulators (SIG) (black bars). Values from duplicate arrays were collapsed and gene-expression profiles were hierarchically clustered. The rightmost color-coded column indicates the five major expression clusters.

(C and D) mRNA expression profiles of candidate (C) and canonical (D) TLR signaling regulators selected for subsequent experiments. The color-coding of the gene names highlights the corresponding expression clusters from the complete matrix from (A).

See also Figure S1 and Table S1.

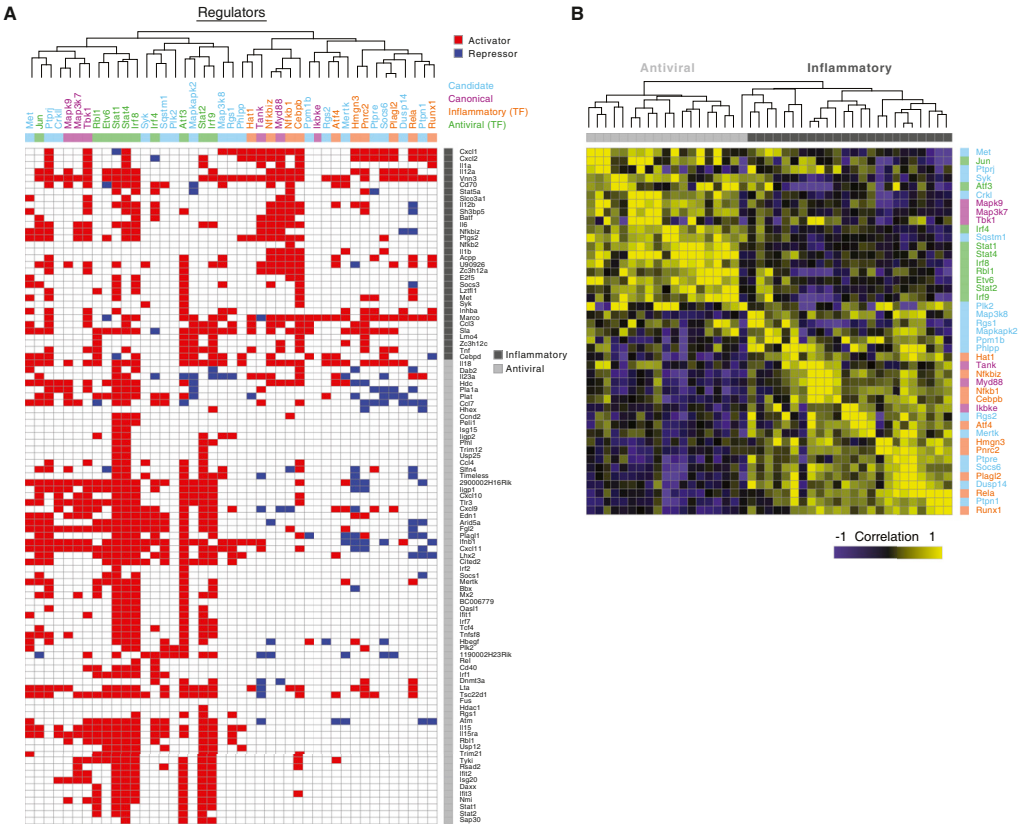

Figure 2. A Perturbation Strategy Assigns Function to Signaling Components within the TLR Pathways

(A) Perturbation profiles of 6 canonical (purple) and 17 candidate (blue) signaling components and 20 core TLR transcriptional regulators belonging to the inflammatory (orange) and the antiviral (green) programs. Shown are the perturbed regulators (columns) and their statistically significant effects (FDR < 0.02) on each of the 118 TLR signature genes (rows). Red: significant activating relation (target gene expression decreased following perturbation); blue: significant repressing relation (target gene expression increased following perturbation); white: no significant effect. The rightmost column categorizes signature genes into antiviral (light gray) and inflammatory (dark gray) programs.

(B) Functional characterization based on similarity of perturbation profiles. Shown is a correlation (Pearson) matrix of the perturbation profiles from (A). Yellow: positive correlation; purple: negative correlation; black: no correlation.

See also Figure S2 and Table S2.

Crkl Modulates JNK-Mediated Antiviral Signaling in the TLR Network

Among the 17 candidate signaling proteins, perturbation of the tyrosine kinase adaptor *Crkl* decreased expression of 13% of the signature genes, especially antiviral ones (Figures 2A and S2D). *Crkl* belongs to several signaling pathways, including early lymphocyte activation (Birge et al., 2009), but has not been associated with the TLR network. *Crkl*'s perturbation profile closely resembled those of known antiviral regulators, most notably Jnk2 (*Mapk9*; Chu et al., 1999) (Figures 2A and 3A). Indeed, when *Crkl*$^{-/-}$ DCs were stimulated with LPS, the expression of antiviral cytokines (*Cxcl10*, *Ifnb1*) was strongly reduced (Figure 3B, left and middle), but that of an inflammatory cytokine (*Cxcl1*) was unaffected (Figure 3B, right).

To test whether *Crkl* is a primary component of the TLR pathway, we determined whether *Crkl* phosphorylation is rapidly modified after TLR signaling initiation. Using SILAC-based (Ong et al., 2002) quantitative phosphoproteomics, we identified and quantified 62 phosphotyrosine (pTyr)-containing peptides from BMDCs stimulated with LPS for 30 min (Figure 3C, Table S3, and Experimental Procedures). Of these 62 phosphopeptides, 7 and 9 were significantly up- or downregulated, respectively

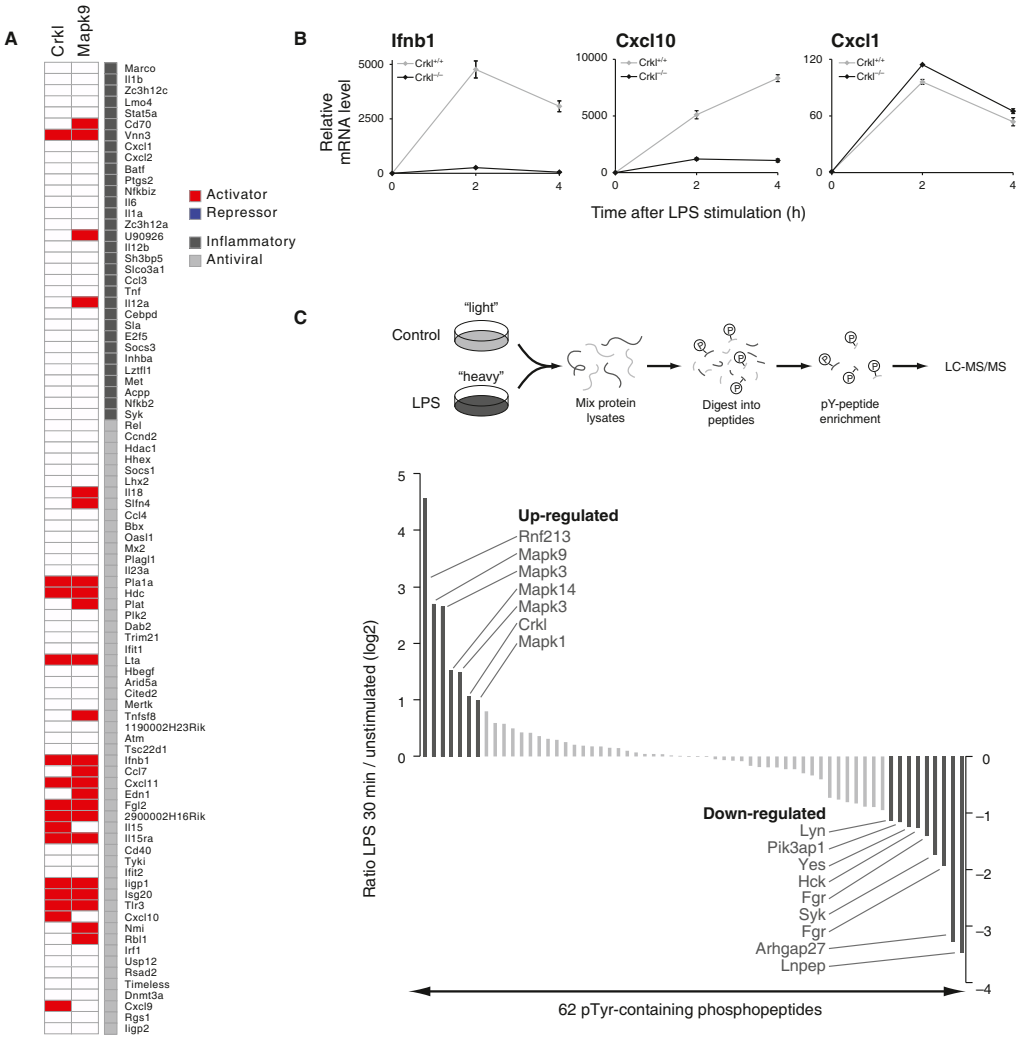

Figure 3. Crkl Adaptor Functions in the Antiviral Arm of TLR4 Signaling

(A) Comparison of Crkl and Mapk9 knockdown profiles. Shown are the effects of Crkl and Mapk9 perturbation (columns) on the 118 signature genes (rows). Data were extracted from Figure 2A.

(B) Inhibition of transcription of antiviral cytokines in Crkl$^{-/-}$ BMDCs. Shown are mRNA levels (qPCR; relative to t = 0) for Ifnb1 (left), Cxcl10 (middle), and Cxcl1 (right) in three replicates per time point. Error bars represent the standard error of the mean (SEM) (n = 3 mice).

(C) Crkl phosphorylation is induced following LPS stimulation. Top: schematic depiction of experimental workflow. From left: Protein lysates from unstimulated (control) and LPS-treated BMDCs grown in "light" and "heavy" SILAC medium were mixed (1:1) and digested into peptides with trypsin before phosphotyrosine (pY) peptide enrichment by immunoprecipitation and LC-MS/MS analysis. Bottom: Shown are the differential phosphorylation levels (log 2 ratios, y axis) of all 62 phosphopeptides identified and quantified by LC-MS/MS (x axis). Black: peptides with more than 2-fold differential expression (left: induced; right: repressed). See also Table S3.

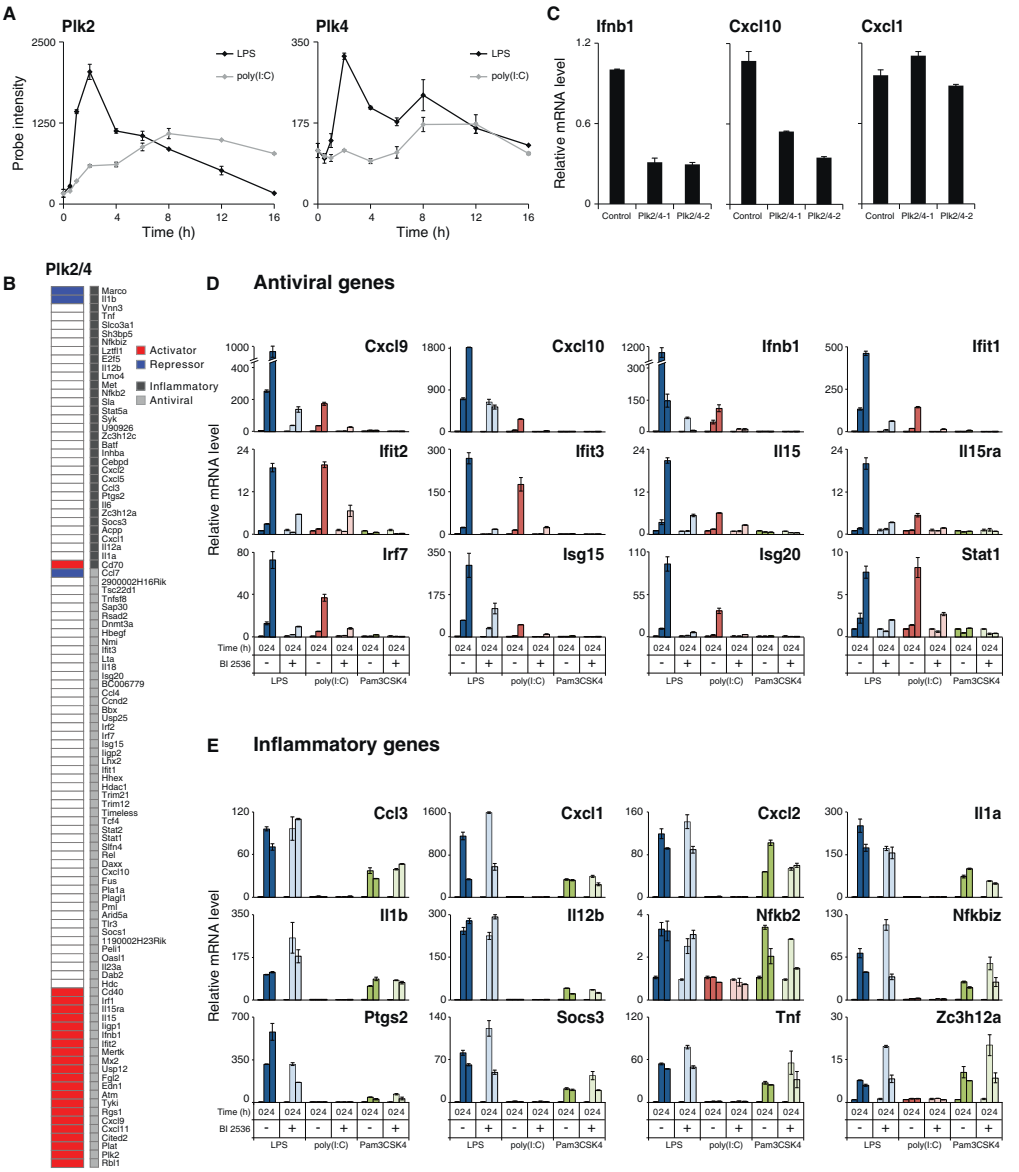

Figure 4. *Plk2* and *4* Regulate the Antiviral Program

(A) Similarity of *Plk2* and *4* mRNA expression profiles. Shown are mRNA levels (from Figure 1B) of *Plk2* (left) and *Plk4* (right) following stimulation with LPS (black) or poly(I:C) (gray). Error bars represent the SEM.

(B) Double knockdown of *Plk2* and *4* represses the antiviral signature. Shown are significant changes in expression of TLR signature genes (rows) following double knockdown of *Plk2* and *4*. Red and blue mark significant hits as in Figure 2, only for genes where the effect was consistent between the two independent combinations of shRNAs.

(Figure 3C and Table S3). A phosphopeptide derived from Crkl (Y132)—one of the top 6 induced phosphopeptides—was induced 2.1-fold (Figure 3C). This indicates that Crkl is likely activated directly downstream of TLR4 signaling.

Several lines of evidence suggest that Crkl acts through Jnk2 (*Mapk9*) signaling. First, the MAP kinase Jnk2 (*Mapk9*) is coregulated at the phosphorylation level with Crkl upon LPS stimulation (Figure 3C). Second, the Crk adaptor family—including Crkl, CrkII, and Crkl—has been shown to modulate Jnk activity in growth factor and IFN signaling (Birge et al., 2009; Hrincius et al., 2010). Third, the perturbation profiles of *Mapk9* and *Crkl* are strikingly similar (Figure 3A). These observations suggest that Crkl modulates Jnk-mediated antiviral signaling in the TLR4 pathway, providing a possible explanation for why the NS1 protein of influenza A virus may target Crkl (Heikkinen et al., 2008; Hrincius et al., 2010).

Polo-like Kinases Are Critical Activators of the Antiviral Program

To discover potential drug targets among our 17 candidates, we next focused on *Plk2*, a well-known cell-cycle regulator and drug target (Strebhardt, 2010). The roles of Plks in nondividing, differentiated cells are poorly defined (Archambault and Glover, 2009; Strebhardt, 2010). We have previously shown that transcriptional regulators of cell-cycle processes (e.g., *Rbl1*, *Rb*, *Myc*, *Jun*, *E2fs*) are co-opted to function in the antiviral responses in DCs (Amit et al., 2009). However, neither knockdown (Figure 2A) nor knockout (Figure S3A) of *Plk2* in BMDCs had any effect on the TLR response. We hypothesized that this could be due to functional redundancy with another Plk, as *Plk4* mRNA was induced in DCs similarly to *Plk2* (Figure 4A), albeit at a lower amplitude (and thus was below our threshold for inclusion in the initial candidate list). Interestingly, functional redundancy between Plk2 and 4 has been suggested to account for the viability of *Plk2*-deficient mice (Strebhardt, 2010), and Plk2 and 4 have been reported to function together in centriole duplication (Chang et al., 2010; Cizmecioglu et al., 2008).

To test our hypothesis, we simultaneously perturbed Plk2 and 4 in BMDCs using two independent mixes of different pairs of shPlk2/shPlk4 (Figure S3B and Experimental Procedures). We observed a significant and specific decrease in the expression of 21 antiviral genes (Figure 4B). For example, the antiviral cytokines *Ifnb1* and *Cxcl10* mRNAs were decreased, whereas the expression of the inflammatory gene *Cxcl1* and almost all inflammatory signature genes remained unaffected (Figure 4C). Two recent reports suggested a role for Plk1 alone as a negative regulator of MAVS (Vitour et al., 2009) and NF-κB (Zhang et al., 2010) in cell lines. However, knockdown of either Plk1 or 3 in BMDCs did not affect the TLR transcriptional response (Figure S3C and Table S2). Notably, BMDC viability was unaffected by lentiviral shRNA transduction targeting Plk1, 2, 3, or 4 individually or Plk2 and 4 together (based on mRNA levels of control genes; Table S2). Thus, in BMDCs, Plk2 and 4, but likely not Plk1 or 3, are critical regulators of antiviral but not cell-cycle pathways.

A Small-Molecule Inhibitor of Plks Represses Antiviral Gene Expression and IRF3 Translocation in DCs

We next targeted Plks in BMDCs using BI 2536, a commercial pan-specific Plk small-molecule inhibitor (Steegmaier et al., 2007). We compared genome-wide mRNA profiles from BMDCs treated with either BI 2536 or DMSO vehicle before stimulation with LPS or poly(I:C) (Experimental Procedures). BI 2536 treatment repressed mostly antiviral gene expression compared to DMSO (99/193 genes in response to poly(I:C), $p < 1 \times 10^{-71}$, hypergeometric test; 67/194 in response to LPS; Table S4). The 311 unique LPS- and/or poly(I:C)-induced genes that are repressed by BI 2536 are significantly enriched for genes related to cytokine signaling (e.g., IL-10, type I IFNs, IL-1), TLR signaling, and DC signaling and for gene ontology (GO) processes related to defense and immune responses (Figure S4A). Consistent with the array data, BI 2536 strongly inhibited the expression of 12 well-studied antiviral genes, whereas inflammatory gene expression remained largely unaffected in DCs stimulated with LPS, poly(I:C), or Pam3CSK4, as measured by qPCR (Figure 4D).

BI 2536 reduced the mRNA levels of *Cxcl10* and *Ifnb1* (by qPCR) and of secreted IFN-β in a dose-dependent manner, whereas *Cxcl1* expression was not significantly affected (Figures S4B and S4C). Importantly, BI 2536 treatment prestimulation impacted neither the viability nor the cell-cycle state of BMDCs (Figures S4D and S4E), suggesting that Plk inhibition does not act through cell-cycle effects. Consistent with our shRNA and BI 2536 perturbations, two other pan-Plk inhibitors—structurally unrelated to BI 2536—also repressed *Ifnb1* and *Cxcl10* expression without affecting *Cxcl1* (Figure S4F). This strongly suggests that the effects induced by these perturbations are due to Plk inhibition and not off-target effects. Furthermore, we observed a similar inhibitory effect of BI 2536 on *Ifnb1* induction in *Ifnar1*$^{-/-}$ and wild-type BMDCs, demonstrating that Plks act directly downstream of TLR activation and not in an autocrine/paracrine feedback loop mediated by IFN receptor signaling (Figure S4G). This is consistent with a recent phosphoproteomic study reporting an enrichment for Plk substrates as early as 15 min after LPS stimulation in macrophages (Weintz et al., 2010).

We next used confocal microscopy to monitor the effect of BI 2536 on the subcellular localization of IRF3, a key antiviral transcription factor. To more effectively deliver the drug, we plated BMDCs on vertical silicon nanowires (Shalek et al., 2010) precoated with BI 2536 prestimulation. Nanowires alone had no effect on the TLR response (Figures 5A and S5A). BI 2536

(C) Double knockdown of *Plk2* and *4* represses antiviral cytokine mRNAs. Shown are expression levels (qPCR) relative to control shRNAs (control) for two antiviral cytokines (*Ifnb1* and *Cxcl10*) and for an inflammatory cytokine (Cxcl1), following LPS stimulation in BMDCs using two independent combinations of shRNAs (*Plk2/4-1*, *Plk2/4-2*). Three replicates for each experiment; error bars are the SEM.

(D and E) BI 2536 specifically abrogates transcription of antiviral genes without affecting inflammatory genes following stimulation with LPS, poly(I:C), or Pam3CSK4. Shown are mRNA levels (qPCR; relative to t = 0) for 12 indicated antiviral (D) and 12 inflammatory (E) genes in BMDCs treated with BI 2536 (1 μM; dark color bars) or DMSO vehicle (light color bars) and stimulated for 0, 2, or 4 hr with LPS (dark and light). Error bars represent the SEM.

See also Figures S3 and S4 and Table S4.

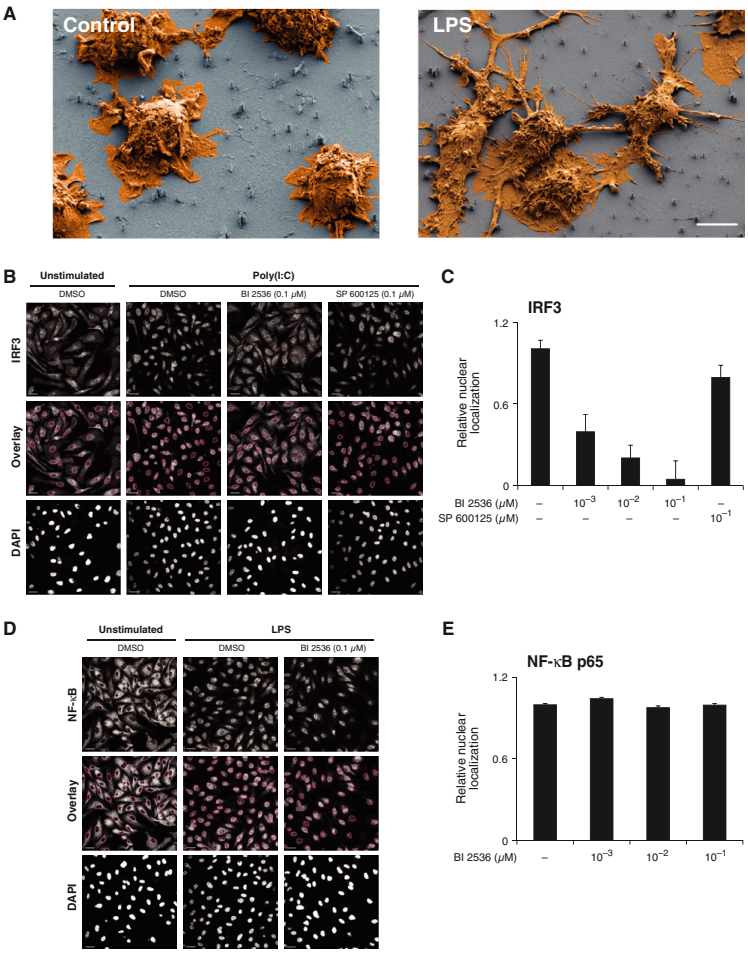

Figure 5. BI 2536-Mediated Plk Inhibition Blocks IRF3 Nuclear Translocation in DCs

(A) DCs on NW undergo normal morphological changes upon LPS stimulation. Shown are electron micrographs of BMDCs plated on bare vertical silicon NW that were left unstimulated (left; control) or stimulated with LPS (right). Scale bars, 5 μm.

(B–E) BI 2536 inhibits IRF3, but not NF-κB p65, nuclear translocation following TLR stimulation. (B and D) Shown are confocal micrographs of BMDCs plated on vertical silicon NW precoated with vehicle control (DMSO; B and D), Plk inhibitor (BI 2536; B and D), or control Jnk inhibitor (SP 600125; B) and stimulated with poly(I:C) for 2 hr (B) or LPS for 30 min (D) (reflecting peak time of nuclear translocation for IRF3 and NF-κB p65, respectively) or left unstimulated (B and D). Cells were analyzed for DAPI (B and D), IRF3 (B), and NF-κB p65 subunit (D) staining. Scale bars, 5 μM. (C and E) Nuclear translocation (from confocal micrographs) of IRF3 (C) and NF-κB p65 (E) was quantified using DAPI staining as a nuclear mask (purple circles; overlay in B and D) to determine the ratio of total versus nuclear fluorescence (y axis) in BMDCs cultured on NW coated with different amounts of BI 2536 or SP 600125 or with vehicle control (DMSO; x axis). Three replicates in each experiment; error bars are the SEM.

See also Figure S5.

inhibited IRF3 nuclear translocation in a dose-dependent manner upon poly(I:C) or LPS stimulation, whereas the control JNK inhibitor SP 600125 had no effect (Figures 5B, 5C, and S5B). On the other hand, BI 2536 did not affect NF-κB p65 localization (Figures 5D and 5E). Notably, IRF3 translocation was also decreased when delivering BI 2536 in solution, but to a lesser extent compared to nanowire-mediated delivery (Figure S5C), highlighting the utility of highly efficient drug delivery methods

to induce homogeneous effects in single-cell assays. Altogether, these results place Plk2 and 4 as critical regulators of the antiviral program, upstream of a major antiviral transcription factor.

Plks Are Essential for Activation of All Well-Established IFN-Inducing Pathways in Conventional and Plasmacytoid DCs

DCs can be broadly categorized into two major subtypes—conventional and plasmacytoid DCs—each relying on distinct mechanisms to induce type I IFNs and antiviral gene expression (Blasius and Beutler, 2010). In conventional DCs (cDCs), antiviral responses are activated through TLR4/3 signaling (via TRIF) or through the cytosolic sensors RIG-I or MDA-5 (via MAVS) (Figure 6A). In plasmacytoid DCs (pDCs; specialized IFN-producing cells), the antiviral response depends solely on endosomal TLR7 and 9 that signal via MYD88 (Figure 6A) (Blasius and Beutler, 2010; Takeuchi and Akira, 2010).

BI 2536 treatment showed that Plks are essential for the viral-sensing pathways in both cDCs and pDCs. In cDCs, BI 2536 inhibited the transcription of antiviral genes (*Ifnb1* and *Cxcl10*) upon infection with each of four viruses: vesicular stomatitis virus (VSV; Figure 6B, top), Sendai virus (SeV; Figure S6A, top), Newcastle disease virus (NDV; Figure S6A, bottom) (all three sensed through RIG-I), and encephalomyocarditis virus (EMCV), sensed through MDA-5 (Figure 6B, bottom and Experimental Procedures). Notably, BI 2536 neither affected the mRNA level of *Cxcl1* (an inflammatory cytokine) in any of the four cases nor affected the response to heat-killed *Listeria monocytogenes*, a natural TLR2 agonist (Figures 6B, S6A, and S6B). In pDCs, BI 2536 treatment nearly abrogated the transcription of mRNAs for the antiviral cytokines *Ifnb1*, *Ifna2*, and *Cxcl10* after stimulation with type A CpG oligonucleotides (CpG-A) or infection with EMCV, sensed by TLR9 and 7, respectively (Figures 6C and S6C and Experimental Procedures). Conversely, in pDCs stimulated with CpG-B—a ligand known to activate inflammatory pathways but not IFN-inducing pathways—BI 2536 treatment decreased *Cxcl10* mRNA, while moderately increasing Cxcl1 mRNA (Figure 6C). Finally, of our 118 signature genes, BI 2536 repressed genes induced by CpG-A alone or by both CpG-A and -B, although having a minor effect, if any, on CpG-B-specific genes in pDCs (Figure 6D and Table S5). These findings may help reveal the poorly characterized molecular determinants of IFN production in pDCs (Reizis et al., 2011) and demonstrate a critical role for Plks across all well-known IFN-inducing pathways.

Plks Are Essential in the Control of Host Antiviral Responses

To assess the impact of Plk inhibition on the outcome of viral infection, we infected primary mouse lung fibroblasts (MLFs) with influenza virus. BI 2536-treated MLFs infected with influenza failed to produce interferon (Figure 6E) and showed elevated replication of both wild-type (PR8) and poorly replicating mutant (ΔNS1) viruses (Figure 6F). The reduced interferon response was not due to drug-induced toxicity (Figure 6G).

Next, we tested the effects of Plk inhibition in virally infected mice. BI 2536 exhibits good tolerability in mice (Steegmaier et al., 2007) and humans (Mross et al., 2008) and is currently in phase II clinical trials as an antitumor agent in several cancers (Strebhardt, 2010). Given its efficacy and safety in vivo, we tested whether BI 2536 would also affect the response to viral infection in animals. In mice infected with VSV, BI 2536 strongly suppressed mRNA production in popliteal lymph nodes for type I IFNs (*Ifnb1*, *Ifna2*) and *Cxcl10* but did not affect *Cxcl1* mRNA induction (all compared to vehicle control; Figures 6H and S6D). Concomitantly, VSV replication in the lymph node rapidly increased as reflected by elevated VSV RNA levels (Figure 6I), comparable to the observed phenotype of VSV-infected *Ifnar1*$^{-/-}$ mice (Iannacone et al., 2010). Because in the VSV model used here type I IFNs are produced by both infected CD169^{+} subcapsular sinus macrophages and pDCs (Iannacone et al., 2010), we cannot distinguish whether Plk inhibition affects macrophages, pDCs, or both. Nevertheless, our results confirm the physiological importance of Plks in the host antiviral response in both ex vivo primary MLFs and in vivo mouse lymph nodes.

Plks Affect the Phosphorylation of Dozens of Proteins Post-LPS Stimulation, including Known and Candidate Antiviral Regulators

We next sought to discover the signaling pathways between Plks and antiviral gene transcription. We used microwestern arrays (MWAs) (Ciaccio et al., 2010) to measure changes in the phosphorylation and protein levels of 20 and 6 TLR pathway proteins, respectively, in BMDCs at each of 12 combinations of four time points (0, 20, 40, 80 min after LPS stimulation) and three perturbations (vehicle control, BI 2536, and negative control JNK inhibitor SP 600125) (Table S6). Although LPS stimulation alone led to the expected changes (e.g., early peak of phosphorylation for ERK1/2, p38, and Mapkapk2 and rapid degradation of IκBα; Figure 7A), BI 2536 surprisingly did not cause any significant changes (Figures 7A, S7A, and S7B). We therefore hypothesized that Plks could affect previously unrecognized regulators of IFN-inducing pathways and/or known regulators with no existing antibodies to specific phosphosites.

Next, we used SILAC-based unbiased phosphoproteomics (Figure 7B, top) (Villén and Gygi, 2008) to compare the levels of phosphotyrosine, -threonine, and -serine peptides following stimulation with LPS (for 30 or 120 min) in BMDCs pretreated with BI 2536 versus those treated with vehicle (DMSO). We identified and quantified 5,061 and 5,997 phosphopeptides after 30 and 120 min, respectively, for a total of 10,236 individual phosphosites (Figure 7B and Table S6). BI 2536 substantially affected the TLR phosphoproteome, leading to a significant ($p < 0.001$) change in the level of 510 phosphopeptides derived from 413 distinct proteins (Figure 7B and Table S6). Further supporting our results, 35% (2489/7018) of the phosphosites we identified were recently reported in mouse bone marrow-derived macrophages treated with LPS (Figure S7C, left) (Weintz et al., 2010), and 483 of our phosphosites were among 1,858 sites (26%) reported in a phosphoproteomic study of LPS signaling in a macrophage cell line (Figure S7C, left) (Sharma et al., 2010). A comparison of the phosphosites of known kinases showed similar overlaps between the three studies (Figure S7C, right).

The Plk-dependent phosphoproteins include several known regulators of antiviral pathways (e.g., Prdm1, Fos, Unc13d) (Crozat et al., 2007; Keller and Maniatis, 1991; Takayanagi et al.,

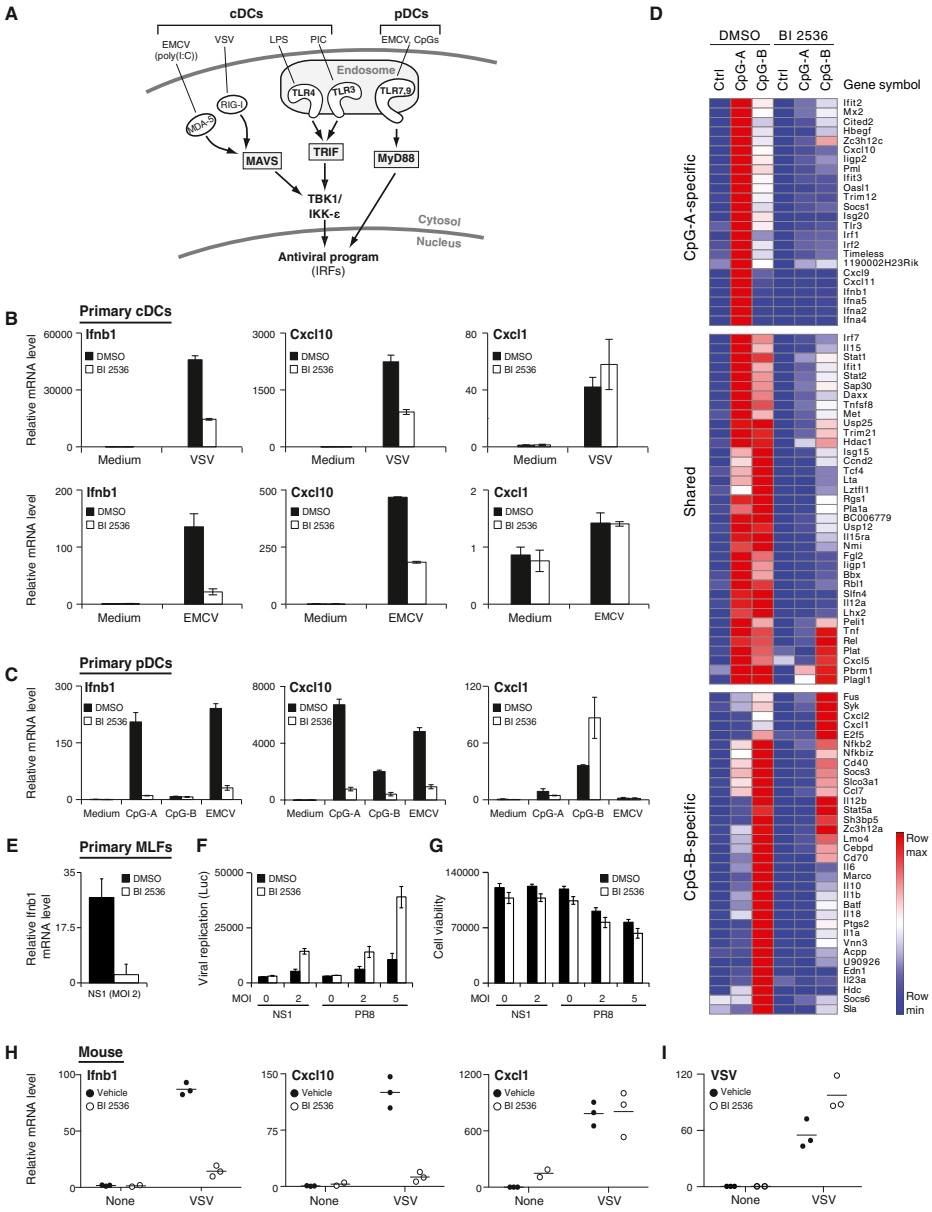

Figure 6. Plks Are Critical in the Induction of Type I Interferons In Vitro and In Vivo

(A) IFN-inducing pathways in cDCs and pDCs.

(B and C) BI 2536 inhibits mRNA levels for antiviral cytokines in response to diverse stimuli in cDCs and pDCs. Shown are Ifnb1, Cxcl10, and Cxcl1 mRNA levels (qPCR; relative to t = 0) in cells treated with BI 2536 (1 μM; white bars) or DMSO vehicle (black bars) in cDCs (B) infected with VSV (multiplicity of infection

2002), as well as many additional protein candidates with no previously known function in viral sensing (Figure 7B and Table S6). Notably, proteins involved in the TBK1/IKK-ε/IRF3 axis were detected and quantified, but their phosphorylation levels were unchanged upon Plk inhibiton (Table S6), consistent with the MWA data. Conversely, Plk inhibition with BI 2536 decreased the phosphorylation levels of cell-cycle regulators of the Jun family of transcriptional regulators (i.e., *Jund*) that we previously found to be co-opted by antiviral pathways (Amit et al., 2009). BI 2536 treatment also decreased the phosphorylation levels of the mitotic kinases Nek6 and Nek7 (Figure 7B). The recent observation that the phosphorylation of Nek6 substrates is increased following LPS stimulation in macrophages (Weintz et al., 2010) indirectly corroborates our finding that Nek6 may be active in TLR signaling. To test the role of these Plk-dependent candidates, we returned to our shRNA perturbation-based approach.

Plk-Dependent Phosphoproteins Affect the Antiviral Response

We perturbed 25 Plk-dependent phosphoproteins (Table S7), using shRNA perturbation in BMDCs followed by qPCR and TLR gene signature measurements. These candidates satisfied three criteria: (1) there was no prior knowledge of their function in viral-sensing pathways; (2) their phosphoprotein levels were consistently up- or downregulated upon BI 2536 treatment (in two independent experiments); and (3) they had detectable mRNA expression and/or differential expression upon stimulation.

Of the 18 phosphoproteins showing efficient knockdown, 11 caused a significant decrease in Ifnb1 mRNA levels with a single shRNA (*Sash1*, *Dock8*, *Nek6*, *Nek7*, *Nfatc2*, and *Ankrd17*; Figure S7D) or with two independent shRNAs (*Tnfaip2*, *Samsn1*, *Arhgap21*, *Mark2*, and *Zc3h14*; Figure S7E). Decrease in *Cxcl10* expression was less prominent, consistent with our previous observations of BI 2536's weaker effect on this cytokine during LPS stimulation (Figures S7D and S7E, far right panels). Each of the 11 Plk-dependent phosphoproteins tested affected at least 9 targets in the 118-gene signature (on average, 39 targets ± 30 SD; Figure 7C), and 9 affected more than 10% of the targets in the TLR gene signature (Figure 7C).

Nine of the 11 Plk-dependent phosphoproteins affected the TLR signature comparably to major antiviral regulators (Figure 7D). For example, the knockdown profiles of *Samsn1*, *Dock8*, and *Sash1* were closely correlated to those of *Stat* and *Irf* family members (Figure 7D), and those of *Tnfaip2* and *Zc3h14* were most correlated to the *Plk2/4* double knockdown. Interestingly, *Tnfaip2*, a protein of unknown molecular function, has been associated with rheumatoid arthritis and autoimmune myocarditis in genome-wide association studies (Wellcome Trust Case Control Consortium, 2007; Kuan et al., 1999). Our findings provide a potential molecular context for this disease association.

DISCUSSION

Using an integrative strategy combining transcriptomics, genetic and chemical perturbations, and unbiased phosphoproteomics, we established a role for Plks in host defense pathways inducing type I IFNs, likely by controlling the phosphorylation and activity of a module of at least 11 components (Figure 7E). Our findings and approach open up several avenues for future investigations.

Consistent with our finding that cell-cycle transcription factors play a role in antiviral responses (Amit et al., 2009), we identified several cell-cycle kinases (Plks, Neks) as important regulators of these responses. Despite extensive studies on the role of Plk1 in mitosis, the functions of its paralogs—Plk2, 3, and 4—are poorly defined (Strebhardt, 2010). Although they are less essential than Plk1 in regulating cell division, their roles in nondividing cells such as neurons are emerging (Archambault and Glover, 2009; Seeburg et al., 2005). Interestingly, silencing of both Plk2 and 4 was required to reveal their importance in antiviral responses, highlighting the necessity of epistasis analysis in studying mammalian signaling networks. Although it is currently not feasible to screen for genetic interactions at a genome-wide scale, it will be interesting to develop innovative approaches to uncover them.

BI 2536 blocked the nuclear translocation of IRF3 without affecting its phosphorylation level (based on MWAs and phosphoproteomics). A similar phenomenon has been reported for NF-kB (Ye et al., 2011). This suggests that IRF3 translocation in our system is likely to be regulated by a mechanism that does not impact phosphorylation.

Furthermore, Plk inhibition suppresses type I IFN production in vivo during viral infection—a finding that has potential clinical implications. Indeed, disease activity in patients with Systemic Lupus Erythematosus (SLE) correlates with IFN expression signatures (Banchereau and Pascual, 2006), and lupus-prone mice exhibit reduced symptoms upon treatment with a dual inhibitor of TLR7 and 9 (Barrat and Coffman, 2008) or deletion

[moi] 1; B, top) or with EMCV (moi 10; B, bottom) and in pDCs (C) stimulated with CpG type A or B or infected with EMCV (moi 10). Three replicates in each experiment; error bars are the SEM.

(D) BI 2536 inhibits the CpG-A response but has little effect on the CpG-B response. Shown are mRNA levels (nCounter) for the 118 TLR signature genes (rows) in pDCs treated with DMSO vehicle or BI 2536 (1 μM) and left untreated (Ctrl) or stimulated with CpG-A or -B for 6 hr (columns). Three clusters of genes are shown: CpG-A-specific (top), CpG-B-specific (bottom), and shared by CpG-A and -B (middle).

(E–G) BI 2536 inhibits IFN-β production in primary MLFs, leading to an increase in viral replication. MLFs treated with BI 2536 (1 μM; white bars) or vehicle control (DMSO; black bars) were infected with influenza ΔNS1 or PR8 strains at indicated mois. Shown are Ifnb1 mRNA levels measured by qPCR (relative to t = 0; E), viral replication as measured by luciferase (Luc) activity in reporter cells (F), and cell viability measured by CellTiter-Glo assay (G). Error bars represent the SEM.

(H and I) BI 2536 inhibits antiviral cytokine mRNA production, while increasing viral replication during in vivo VSV infection. Shown are *Ifnb1*, *Cxcl10*, and *Cxcl1* mRNA (H) and VSV viral RNA (I) levels (qPCR; relative to uninfected animals) from popliteal lymph nodes of mice injected with BI 2536 (white circles) or DMSO vehicle (black circles) prior to and during the course of infection with VSV (intra-footpad). Nodes were harvested 6 hr post-infection. Each circle represents one animal (n = 3). Data are representative of three independent experiments for each condition.

See also Figure S6 and Table S5.

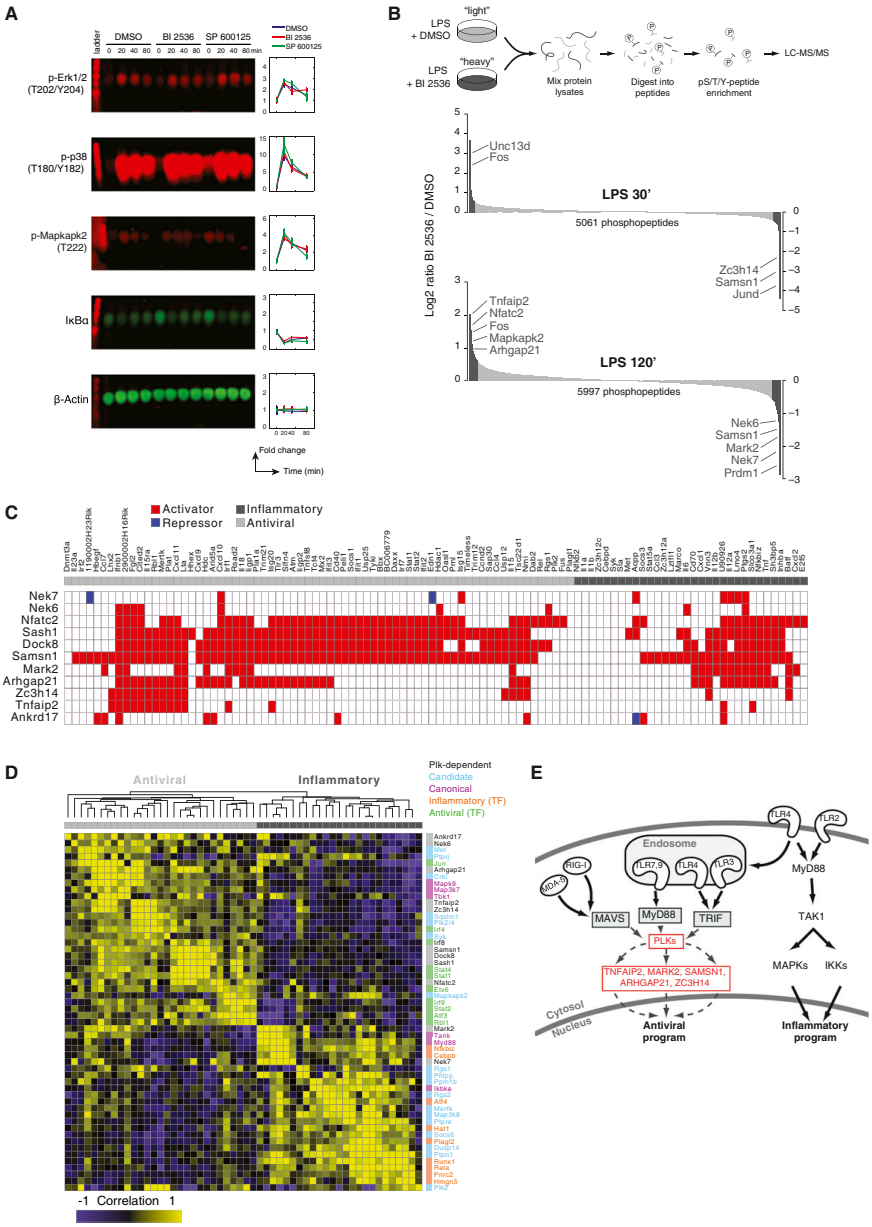

Figure 7. Unbiased Phosphoproteomics Identifies a Plk-Dependent Antiviral Pathway

(A) BI 2536 does not affect phosphorylation and protein levels of known TLR signaling nodes. Shown are representative MWA (see Experimental Procedures) blots (left) obtained from analyzing lysates from BMDCs pretreated with DMSO, BI 2536 (1 μM), or SP 600125 (5 μM) and stimulated with LPS for 0, 20, 40, and

of the IFN receptor (Santiago-Raber et al., 2003). Thus, testing the effect of BI 2536 on a mouse model of lupus will be key to assess the potential therapeutic implications of Plk inhibition for SLE.

Our approach may be applicable for characterizing the functions of genes reported in genome-wide association studies (e.g., *Tnfaip2*), for uncovering potential therapeutic targets (e.g., Plks), and for repurposing existing small molecules in new physiological contexts (e.g., using the cancer drug BI 2536 to repress innate immune responses). The vast public compendia of microarray data could serve as starting points for identification of relevant signaling components in diverse biological systems, followed by perturbations and signature measurements. Nevertheless, because the mRNAs corresponding to many pathway components do not change upon pathway activation, our approach is far from exhaustive. Combination of our perturbation-based approach with large-scale biochemical measurements (e.g., posttranslational modifications, protein-protein interactions) will lead to more comprehensive, integrated maps of signaling and transcriptional networks.

EXPERIMENTAL PROCEDURES

Cells and Mouse Strains

BMDCs were generated from 6- to 8-week-old female C57BL/6J mice, *Crkl* mutant mice (Jackson Laboratories), *Plk2*$^{-/-}$ mice (Elan Pharmaceuticals), or *Ifnar1*$^{-/-}$ mice (gift from K. Fitzgerald). Primary MLFs were from C57BL/6J mice.

Viruses

SeV strain Cantell and EMCV strain EMC (ATCC), NDV strain Hitchner B1 (gift from A. Garcia-Sastre), and VSV strain Indiana (U. von Andrian) were used for infections. Influenza A virus strain A/PR/8/34 and ΔNS1 were grown in Vero cells, and virus titers from MLF supernatants were quantified using 293T cells transfected with a vRNA luciferase reporter plasmid.

Reagents

TLR ligands were from Invivogen (Pam3CSK4, ultra-pure *E. coli* K12 LPS, ODN 1585 CpG type A, and ODN 1668 CpG type B) and Enzo Life Sciences (poly(I:C)). Heat-killed *Listeria monocytogenes* (HKLM) was from Invivogen. Polo-like kinase inhibitors were from Selleck (BI 2536), Sigma (GW843682X), and Chembridge (Poloxipan). SP 600125 (Jnk inhibitor) was from Enzo Life Sciences.

mRNA Isolation, qPCR, and Microarrays

Total or poly(A)$^+$ RNA was extracted and reverse transcribed prior to qPCR analysis with SYBR Green (Roche) in triplicate with GAPDH for normalization. For microarray analysis, Affymetrix Mouse Genome 430A 2.0 Array was used.

shRNA Knockdowns

High-titer lentiviruses expressing shRNAs were obtained from The Broad RNAi Platform and used to infect BMDCs as previously described (Amit et al., 2009).

mRNA Measurements on nCounter

5×10^4 BMDCs were lysed in RLT buffer (QIAGEN) with 1% β-ME. Ten percent of the lysate was used for mRNA counting using the nCounter Digital Analyzer (NanoString) and a previously generated CodeSet of 118 genes (Amit et al., 2009). To score target genes whose expression is significantly affected by shRNA perturbations, we used a fold threshold corresponding to a false discovery rate (FDR) of 2%. Heatmaps and distance matrix analyses were generated using the Gene-E software (http://www.broadinstitute.org/cancer/software/GENE-E/).

Detection of Regulated Signaling Genes

We identified differentially regulated signaling components (i.e., kinases, phosphatases, and signaling adaptors or scaffolds) based on probesets reproducibly displaying at least 1.7-fold up- or downregulation in at least one time point, compared to unstimulated controls, using our previously published microarray dataset (NCBI GEO GSE17721, Amit et al., 2009).

Nanowire-Mediated Drug Delivery and Microscopy

BMDCs were plated on top of etched silicon nanowires (Si NWs) coated with small molecules. After 24 hr, cells were stimulated and processed for immunofluorescence analysis by confocal microscopy.

VSV Infection Model

Eight-week-old C57BL/6 male mice received 500 μg of BI 2536 (or vehicle) intravenously and 50 μg into the footpad 3 hr before and 2 hr after infection with 10^6 pfu of VSV into the footpad. Mice were sacrificed 6 hr post-infection, and the draining popliteal lymph nodes were harvested in RNAlater solution (Ambion) before subsequent RNA extraction and qPCR analysis.

Microwestern Arrays

The MWA method previously described (Ciaccio et al., 2010) was modified to accommodate a larger number of lysates.

Phosphotyrosine and Global Phospopeptide Analysis

Tyrosine-phosphorylated peptides from BMDC lysates were prepared using a PhosphoScan Kit (Cell Signaling Technology) and analyzed by data-dependent LC-MS/MS using a Thermo LTQ-Orbitrap. Quantitative analysis of

80 min. Blots were analyzed using indicated antibodies (leftmost), and fold change in fluorescence signals was quantified relative to t = 0 (right). Error bars are the SEM of triplicate MWA blots.

(B) BI 2536 affects protein phosphorylation levels during LPS stimulation. Top: Schematic depiction of experimental workflow. From left to right: LPS-stimulated BMDCs cultured in "heavy" or "light" SILAC medium were pretreated with BI 2536 (1 μM) or DMSO, respectively. Protein lysates were mixed (1:1) and digested into peptides with trypsin, before phosphoserine, -threonine, and -tyrosine (pS/T/Y) peptide enrichment and LC-MS/MS analysis. Bottom: Shown are the differential phosphorylation levels (average $\log_2$ ratios of two independent experiments; y axis) of all 5,061 and 5,997 phosphopeptides, respectively, identified and quantified by LC-MS/MS (x axis) at 30 min (top) and 120 min (bottom) post-LPS stimulation. Dark gray: phosphopeptides with a significant change ($P_{unadjusted} < 0.001$ for both time points; $FDR_{30min} = 0.05$; $FDR_{120min} = 0.03$; left: induced; right: repressed). Average ratios from phosphopeptides identified and quantified in two independent experiments are depicted.

(C) Eleven Plk-dependent phosphoproteins significantly affect the expression of TLR signature genes. Shown are significant changes in expression of the TLR signature genes (rows) following knockdown of each of the 11 phosphoproteins (columns). Cells were stimulated with LPS for 6 hr. Red and blue mark significant hits (as presented in Figure 2) and are shown only for genes where the effect was consistent between two independent experiments.

(D) Functional characterization based on similarity of perturbation profiles. Shown is a correlation (Pearson) matrix of the perturbation profiles from (C) (gray) and those from Figure 2B including canonical (purple) and candidate (blue) signaling components as well as core antiviral (green) and inflammatory (orange) transcriptional regulators. Yellow: positive correlation; purple: negative correlation; black: no correlation.

(E) A Plk-dependent pathway in antiviral sensing. Shown is a diagram of a model of the Plk-dependent pathway of IFN induction in innate immunity. Out of the 11 Plk-dependent proteins described in (C) and (D), only the 5 showing a phenotype with two independent shRNAs are depicted.

See also Figure S7 and Tables S6 and S7.

serine-, threonine-, and tyrosine-phosphorylated peptides was performed using SCX/IMAC as described (Villén and Gygi, 2008) with some modifications. Peptide samples were analyzed on a LTQ-Orbitrap Velos (Thermo Fisher Scientific). To identify and quantify peptides, mass spectra were processed with Spectrum Mill software package (Agilent Technologies) v4.0b, including in-house developed features for SILAC quantitation and phosphosite localization, and with MaxQuant (v1.0.13.13) (Cox and Mann, 2008) and Mascot search engine (v2.2.0, Matrix Science).

ACCESSION NUMBERS

Complete microarray datasets are available in the NCBI Gene Expression Omnibus (accession number GSE28520). Proteomics raw data are in the the Tranche data repository (https://proteomecommons.org/tranche/, hash: HTWY5ZeSLM1hyYEyfEiJREkgLXs6BZxCczuixy9XjULsync5HCkXx/8gB7nZKpGocwOnt8vOk/Q3cpbPh/ycD/2LT0AAAAAAAAAuEg = =, and passphrase: SpSTB6vceSUKeNqefq59).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, seven figures, and seven tables and can be found with this article online at doi:10.1016/j.cell.2011.10.022.

ACKNOWLEDGMENTS

We thank the members of the Hacohen and Regev laboratories, O. Takeuchi, B. Beutler, D. Mathis, P. Anderson, L. Glimcher, and D. Bartel for valuable discussions and comments; L. Gaffney and L. Solomon for assistance with figures and artwork; and S. Gupta and the Broad Genetic Analysis Platform for microarray processing. This work was supported by NIH grant U54 AI057159, the NIH New Innovator Award DP2 OD002230 (N.H.), and NIH P50 HG006193 (A.R., N.H.); a Ph.D fellowship from the Boehringer Ingelheim Fonds (N.C.); the Human Frontier Science Program Organization and a Claire and Emanuel G. Rosenblatt Award from the American Physicians Fellowship for Medicine in Israel (I.A.); NIH Pioneer Awards (A.R., H.P.); and a Career Award at the Scientific Interface from the Burroughs Wellcome Fund, the Sloan Foundation, and HHMI (A.R.). A.R. is an investigator of the Merkin Foundation for Stem Cell Research at the Broad Institute. This work was partially supported by a pilot award from the NIH Chicago Center for Systems Biology P50 GM081892-03 to R.B.J. and an award from the American Cancer Society Illinois Division to R.B.J.

Received: February 28, 2011
Revised: April 27, 2011
Accepted: October 17, 2011
Published: November 10, 2011

REFERENCES

Amit, I., Citri, A., Shay, T., Lu, Y., Katz, M., Zhang, F., Tarcic, G., Siwak, D., Lahad, J., Jacob-Hirsch, J., et al. (2007). A module of negative feedback regulators defines growth factor signaling. Nat. Genet. *39*, 503–512.

Amit, I., Garber, M., Chevrier, N., Leite, A.P., Donner, Y., Eisenhaure, T., Guttman, M., Grenier, J.K., Li, W., Zuk, O., et al. (2009). Unbiased reconstruction of a mammalian transcriptional network mediating pathogen responses. Science *326*, 257–263.

Archambault, V., and Glover, D.M. (2009). Polo-like kinases: conservation and divergence in their functions and regulation. Nat. Rev. Mol. Cell Biol. *10*, 265–275.

Banchereau, J., and Pascual, V. (2006). Type I interferon in systemic lupus erythematosus and other autoimmune diseases. Immunity *25*, 383–392.

Barbie, D.A., Tamayo, P., Boehm, J.S., Kim, S.Y., Moody, S.E., Dunn, I.F., Schinzel, A.C., Sandy, P., Meylan, E., Scholl, C., et al. (2009). Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature *462*, 108–112.

Barrat, F.J., and Coffman, R.L. (2008). Development of TLR inhibitors for the treatment of autoimmune diseases. Immunol. Rev. *223*, 271–283.

Birge, R.B., Kalodimos, C., Inagaki, F., and Tanaka, S. (2009). Crk and CrkL adaptor proteins: networks for physiological and pathological signaling. Cell Commun. Signal. *7*, 13.

Blasius, A.L., and Beutler, B. (2010). Intracellular toll-like receptors. Immunity *32*, 305–315.

Chang, J., Cizmecioglu, O., Hoffmann, I., and Rhee, K. (2010). PLK2 phosphorylation is critical for CPAP function in procentriole formation during the centrosome cycle. EMBO J. *29*, 2395–2406.

Chien, Y., Kim, S., Bumeister, R., Loo, Y.M., Kwon, S.W., Johnson, C.L., Balakireva, M.G., Romeo, Y., Kopelovich, L., Gale, M., Jr., et al. (2006). RalB GTPase-mediated activation of the IkappaB family kinase TBK1 couples innate immune signaling to tumor cell survival. Cell *127*, 157–170.

Chu, W.M., Ostertag, D., Li, Z.W., Chang, L., Chen, Y., Hu, Y., Williams, B., Perrault, J., and Karin, M. (1999). JNK2 and IKKbeta are required for activating the innate response to viral infection. Immunity *11*, 721–731.

Ciaccio, M.F., Wagner, J.P., Chuu, C.P., Lauffenburger, D.A., and Jones, R.B. (2010). Systems analysis of EGF receptor signaling dynamics with microwestern arrays. Nat. Methods *7*, 148–155.

Cizmecioglu, O., Warnke, S., Arnold, M., Duensing, S., and Hoffmann, I. (2008). Plk2 regulated centriole duplication is dependent on its localization to the centrioles and a functional polo-box domain. Cell Cycle *7*, 3548–3555.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. *26*, 1367–1372.

Crozat, K., Hoebe, K., Ugolini, S., Hong, N.A., Janssen, E., Rutschmann, S., Mudd, S., Sovath, S., Vivier, E., and Beutler, B. (2007). Jinx, an MCMV susceptibility phenotype caused by disruption of Unc13d: a mouse model of type 3 familial hemophagocytic lymphohistiocytosis. J. Exp. Med. *204*, 853–863.

Doyle, S., Vaidya, S., O'Connell, R., Dadgostar, H., Dempsey, P., Wu, T., Rao, G., Sun, R., Haberland, M., Modlin, R., and Cheng, G. (2002). IRF3 mediates a TLR3/TLR4-specific antiviral gene program. Immunity *17*, 251–263.

Fraser, I.D., and Germain, R.N. (2009). Navigating the network: signaling cross-talk in hematopoietic cells. Nat. Immunol. *10*, 327–331.

Freeman, M. (2000). Feedback control of intercellular signalling in development. Nature *408*, 313–319.

Heikkinen, L.S., Kazlauskas, A., Melén, K., Wagner, R., Ziegler, T., Julkunen, I., and Saksela, K. (2008). Avian and 1918 Spanish influenza a virus NS1 proteins bind to Crk/CrkL Src homology 3 domains to activate host cell signaling. J. Biol. Chem. *283*, 5719–5727.

Hennessy, E.J., Parker, A.E., and O'Neill, L.A. (2010). Targeting Toll-like receptors: emerging therapeutics? Nat. Rev. Drug Discov. *9*, 293–307.

Hrincius, E.R., Wixler, V., Wolff, T., Wagner, R., Ludwig, S., and Ehrhardt, C. (2010). CRK adaptor protein expression is required for efficient replication of avian influenza A viruses and controls JNK-mediated apoptotic responses. Cell. Microbiol. *12*, 831–843.

Iannacone, M., Moseman, E.A., Tonti, E., Bosurgi, L., Junt, T., Henrickson, S.E., Whelan, S.P., Guidotti, L.G., and von Andrian, U.H. (2010). Subcapsular sinus macrophages prevent CNS invasion on peripheral infection with a neurotropic virus. Nature *465*, 1079–1083.

Kawagoe, T., Takeuchi, O., Takabatake, Y., Kato, H., Isaka, Y., Tsujimura, T., and Akira, S. (2009). TANK is a negative regulator of Toll-like receptor signaling and is critical for the prevention of autoimmune nephritis. Nat. Immunol. *10*, 965–972.

Keller, A.D., and Maniatis, T. (1991). Identification and characterization of a novel repressor of beta-interferon gene expression. Genes Dev. *5*, 868–879.

Kuan, A.P., Chamberlain, W., Malkiel, S., Lieu, H.D., Factor, S.M., Diamond, B., and Kotzin, B.L. (1999). Genetic control of autoimmune myocarditis mediated by myosin-specific antibodies. Immunogenetics *49*, 79–85.

Matsui, K., Kumagai, Y., Kato, H., Sato, S., Kawagoe, T., Uematsu, S., Takeuchi, O., and Akira, S. (2006). Cutting edge: Role of TANK-binding kinase 1

and inducible IkappaB kinase in IFN responses against viruses in innate immune cells. J. Immunol. *177*, 5785–5789.

Mross, K., Frost, A., Steinbild, S., Hedbom, S., Rentschler, J., Kaiser, R., Rouyrre, N., Trommeshauser, D., Hoesl, C.E., and Munzert, G. (2008). Phase I dose escalation and pharmacokinetic study of BI 2536, a novel Polo-like kinase 1 inhibitor, in patients with advanced solid tumors. J. Clin. Oncol. *26*, 5511–5517.

Ong, S.E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A., and Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol. Cell. Proteomics *1*, 376–386.

Reizis, B., Bunin, A., Ghosh, H.S., Lewis, K.L., and Sisirak, V. (2011). Plasmacytoid dendritic cells: recent progress and open questions. Annu. Rev. Immunol. *29*, 163–183.

Rothlin, C.V., Ghosh, S., Zuniga, E.I., Oldstone, M.B., and Lemke, G. (2007). TAM receptors are pleiotropic inhibitors of the innate immune response. Cell *131*, 1124–1136.

Santiago-Raber, M.L., Baccala, R., Haraldsson, K.M., Choubey, D., Stewart, T.A., Kono, D.H., and Theofilopoulos, A.N. (2003). Type-I interferon receptor deficiency reduces lupus-like disease in NZB mice. J. Exp. Med. *197*, 777–788.

Seeburg, D.P., Pak, D., and Sheng, M. (2005). Polo-like kinases in the nervous system. Oncogene *24*, 292–298.

Shalek, A.K., Robinson, J.T., Karp, E.S., Lee, J.S., Ahn, D.R., Yoon, M.H., Sutton, A., Jorgolli, M., Gertner, R.S., Gujral, T.S., et al. (2010). Vertical silicon nanowires as a universal platform for delivering biomolecules into living cells. Proc. Natl. Acad. Sci. USA *107*, 1870–1875.

Sharma, K., Kumar, C., Kéri, G., Breitkopf, S.B., Oppermann, F.S., and Daub, H. (2010). Quantitative analysis of kinase-proximal signaling in lipopolysaccharide-induced innate immune response. J. Proteome Res. *9*, 2539–2549.

Strebhardt, K. (2010). Multifaceted polo-like kinases: drug targets and antitargets for cancer therapy. Nat. Rev. Drug Discov. *9*, 643–660.

Takayanagi, H., Kim, S., Matsuo, K., Suzuki, H., Suzuki, T., Sato, K., Yokochi, T., Oda, H., Nakamura, K., Ida, N., et al. (2002). RANKL maintains bone homeostasis through c-Fos-dependent induction of interferon-beta. Nature *416*, 744–749.

Takeuchi, O., and Akira, S. (2010). Pattern recognition receptors and inflammation. Cell *140*, 805–820.

Villén, J., and Gygi, S.P. (2008). The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry. Nat. Protoc. *3*, 1630–1638.

Vitour, D., Dabo, S., Ahmadi Pour, M., Vilasco, M., Vidalain, P.O., Jacob, Y., Mezel-Lemoine, M., Paz, S., Arguello, M., Lin, R., et al. (2009). Polo-like kinase 1 (PLK1) regulates interferon (IFN) induction by MAVS. J. Biol. Chem. *284*, 21797–21809.

Wang, L., Gordon, R.A., Huynh, L., Su, X., Park Min, K.H., Han, J., Arthur, J.S., Kalliolias, G.D., and Ivashkiv, L.B. (2010). Indirect inhibition of Toll-like receptor and type I interferon responses by ITAM-coupled receptors and integrins. Immunity *32*, 518–530.

Weintz, G., Olsen, J.V., Frühauf, K., Niedzielska, M., Amit, I., Jantsch, J., Mages, J., Frech, C., Dölken, L., Mann, M., and Lang, R. (2010). The phosphoproteome of toll-like receptor-activated macrophages. Mol. Syst. Biol. *6*, 371.

Wellcome Trust Case Control Consortium. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature *447*, 661–678.

Xiao, N., Eidenschenk, C., Krebs, P., Brandl, K., Blasius, A.L., Xia, Y., Khovananth, K., Smart, N.G., and Beutler, B. (2009). The Tpl2 mutation Sluggish impairs type I IFN production and increases susceptibility to group B streptococcal disease. J. Immunol. *183*, 7975–7983.

Ye, J., Chen, S., and Maniatis, T. (2011). Cardiac glycosides are potent inhibitors of interferon-β gene expression. Nat. Chem. Biol. *7*, 25–33.

Zhang, W., Wang, J., Zhang, Y., Yuan, Y., Guan, W., Jin, C., Chen, H., Wang, X., Yang, X., and He, F. (2010). The scaffold protein TANK/I-TRAF inhibits NF-kappaB activation by recruiting polo-like kinase 1. Mol. Biol. Cell *21*, 2500–2513.

EXTENDED EXPERIMENTAL PROCEDURES

Preparation of Dendritic Cells

BMDCs were generated from 6- to 8-week-old female C57BL/6J mice (Jackson Laboratories). Bone marrow cells were collected from femora and tibiae and plated at 10^6 cells/ml on nontissue culture treated petri dishes in RPMI-1640 medium (GIBCO), supplemented with 10% FBS, L-glutamine, penicillin/streptomycin, MEM nonessential amino acids, HEPES, sodium pyruvate, β-mercaptoethanol, and murine GM-CSF (15 ng/ml; Peprotech) or human Flt3L (100 ng/ml; Peprotech). GM-CSF-derived BMDCs were used directly for all RNAi experiments. For all other experiments, floating cells from GM-CSF cultures were sorted at day 5 by MACS using the CD11c (N418) MicroBeads kit (Miltenyi Biotec). Sorted $CD11c^+$ cells were used as GM-CSF-derived BMDCs, and plated at 10^6 cells/ml and stimulated at 16 hr post sorting. For Flt3L culture, floating cells were harvested at days 6–8 and used as Flt3L-derived BMDCs by plating them at 10^6 cells/ml and stimulating 16 hr later.

For SILAC experiments, GM-CSF-derived BMDCs were grown in media containing either normal L-arginine (Arg0) and L-lysine (Lys0) (Sigma) or L-arginine 13C6-15N4 (Arg10) and L-lysine 13C6-15N2 (Lys8) (Sigma Isotec). Concentrations for L-arginine and L-lysine were 42 mg/l and 40 mg/l, respectively. The cell culture media, RPMI-1640 deficient in L-arginine and L-lysine, was a custom media preparation from Caisson Laboratories (North Logan, UT) and dialyzed serum was obtained from SAFC-Sigma. We followed all standard SILAC media preparation and labeling steps as previously described (Ong and Mann, 2006).

Preparation of Primary Lung Fibroblasts

MLFs were derived from lung tissue from 6- to 8-week-old female C57BL/6J mice (Jackson Laboratories). MLFs were isolated as previously described (Tager et al., 2004). Briefly, lungs were digested for 45 min at 37°C in collagenase and DNase I, filtered, washed, and cultured in DMEM supplemented with 15% FBS. Cells were used for experiments between passages 2 and 5.

Genetically Modified Mice

Bone marrow from *Plk2*$^{-/-}$ mice and their wild-type littermates were obtained from Elan Pharmaceuticals (Inglis et al., 2009). *Ifnar1*$^{-/-}$ mice on a C57BL/6J background were a gift from Kate Fitzgerald (originally from Jonathan Sprent based on Muller et al., 1994). Heterozygous *Crkl*$^{+/-}$ mice on a C57BL/6J background were obtained from the Jackson Laboratory. *Crkl*$^{+/-}$ C57BL/6J mice were crossed to wild-type Black Swiss mice from Taconic, as *Crkl*$^{-/-}$ mice on a pure C57BL/6J genetic background have been reported to be embryonic lethal (Guris et al., 2001; Hemmeryckx et al., 2002). Heterozygous *Crkl*$^{+/-}$ offspring were backcrossed to *Crkl*$^{+/-}$ C57BL/6J mice to obtain *Crkl*$^{-/-}$ mice. Mice were kept in a specific pathogen-free facility at MIT. Animal procedures were in accordance with National Institutes of Health Guidelines on animal care and use and were approved by the MIT Committee on Animal Care (Protocol #0609-058-12).

Viruses

SeV, strain Cantell, and EMCV, strain EMC, were from ATCC. NDV, strain Hitchner B1 was from Aldolfo Garcia-Sastre (Mount Sinai School of Medicine), and VSV, strain Indiana was from Ulrich von Andrian (Harvard Medical School). Influenza A virus strain A/PR/8/34 and ΔNS1 were grown in Vero cells (which allow efficient growth of the ΔNS1 virus) in serum-free DMEM supplemented with 10% BSA and 1 mg/ml TPCK trypsin. Viral titers were determined by standard MDCK plaque assay. To measure the amount of VSV RNA present in infected tissues, we used previously reported qPCR primers: VSV forward 5′-TGATACAGTACAATTATTTTGGGAC-3′, and VSV reverse 5′-GAGACTTTCTGTTACGGGATCTGG-3′ (Hole et al., 2006). Viruses were handled according to CDC and NIH guidelines with protocols approved by the Broad Institutional Biosafety Committee.

Reagents

TLR ligands were from Invivogen (Pam3CSK4, ultra-pure *E. coli* K12 LPS, ODN 1585 CpG type A, and ODN 1668 CpG type B) and Enzo Life Sciences (poly(I:C)), and were used at the following concentrations: Pam3CSK4 (250 ng/ml), poly(I:C) (10 μg/ml), LPS (100 ng/ml), CpG-A (10 μg/ml), CpG-B (10 μg/ml). Heat-killed *Listeria monocytogenes* (HKLM) was from Invivogen. Polo-like kinase inhibitors were from Selleck (BI 2536; Steegmaier et al., 2007), Sigma (GW843682X, also known as compound 1 and GSK461364; Lansing et al., 2007), and Chembridge (Poloxipan; Reindl et al., 2009). SP 600125 (Jnk inhibitor) was from Enzo Life Sciences. Image-iT FX Signal Enhancer, DAPI, and Alexa Fluor Labeled Secondary Antibodies were obtained from Invitrogen. For immunofluorescence, antibodies against IRF3 (4302S) and NF-κB P65 (4764S) were obtained from Cell Signaling Technology. For cell viability assays, Alamar Blue was from Invitrogen and CellTiter-Glo from Promega.

Virus Titering of MLF Supernatant

293T cells were seeded and transfected with a vRNA luciferase reporter plasmid as previously described (Shapira et al., 2009). Briefly, at 24 hr post-transfection, 10^4 transfected reporter cells were reseeded in white Costar plates. Supernatants from influenza-infected MLFs were added to reporter cells and incubated for 24 hr. Reporter activity was measured with firefly luciferase substrate (Steady-Glo, Promega). Luminescence activity was quantified with the Envision Multilabel Reader (Perkin Elmer).

Cell-Cycle Analysis

Cells were fixed in ethanol, washed, and stained for 30 min at room temperature (RT) with propidium iodide (100 μg/ml) prepared in PBS (calcium- and magnesium-free) supplemented with RNase A (2 mg/ml; Novagen) and Triton X-100 (0.1%). Samples were analyzed for DNA content using an Accuri C6 flow cytometer (Accuri) and data was processed using the FlowJo software (Treestar).

ELISA

Cell culture supernatants were assayed using a sandwich ELISA kit for mouse IFN-β (PBL Biomedical Laboratories).

mRNA Isolation

Total RNA was extracted with QIAzol reagent following the miRNeasy kit's procedure (QIAGEN), and sample quality was tested on a 2100 Bioanalyzer (Agilent). RNA was reverse transcribed with the High Capacity cDNA Reverse Transcription kit (Applied Biosystems). For experiments with more than 12 samples, we harvested Poly(A)$^+$ RNA in 96- or 384-well plates with the Turbocapture mRNA kit (QIAGEN) and reverse transcribed with the Sensiscript RT kit (QIAGEN).

qPCR Measurements

Real-time quantitative PCR reactions were performed on the LightCycler 480 system (Roche) with FastStart Universal SYBR Green Master Mix (Roche). Every reaction was run in triplicate and GAPDH levels were used as an endogenous control for normalization.

shRNA Knockdowns

High-titer lentiviruses encoding shRNAs targeting genes of interest were obtained from The RNAi Consortium (TRC; Broad Institute, Cambridge, MA, USA) (Moffat et al., 2006). Bone marrow cells were infected with lentiviruses as described (Amit et al., 2009). For each gene of interest, we tested five shRNAs for knockdown efficiency using qPCR of the target gene. We selected shRNAs with >75% knockdown efficacy. For combinatorial knockdown, two independent mixtures of two lentiviruses encoding validated shRNAs against *Plk2* and *4*, respectively, were used to infect bone marrow cells (two *Plk2*- and two *Plk4*-specific shRNAs were used to generate these mixtures). Lentivirus-infected cells were composed of ~90% CD11c$^+$ cells, which was comparable to sorted BMDCs and to our previous observations (Amit et al., 2009).

mRNA Measurements on nCounter

Details on the nCounter system are presented in full in (Geiss et al., 2008). We used a custom CodeSet constructed to detect a total of 128 genes (including 10 control genes whose expression remain unaffected by TLR stimulation) selected by the GeneSelector algorithm (Amit et al., 2009) as described below. 5 × 10^4 BMDCs were lysed in RLT buffer (QIAGEN) supplemented with 1% β-mercaptoethanol. Ten percent of the lysate was hybridized for 16 hr with the CodeSet and loaded into the nCounter prep station followed by quantification using the nCounter Digital Analyzer following the manufacturer's instructions.

Custom Nanostring CodeSet Construction using the GeneSelector Algorithm

We used the CodeSet that we previously used and described in Amit et al., 2009. Briefly, to choose a set of genes that will capture as much as possible of the information on the expression of all genes, we used an information-theoretic approach. We modeled the expression levels X given the experimental condition C with a naive Bayes model where the expression of gene i under condition c follows a normal distribution $X_i|C=c \sim N(\mu_{ic}, \sigma_i^2)$. In this model, the expression levels of all genes depend on the experimental condition C, and we selected genes that are highly informative about C. Formally, for a set of genes Y we used the conditional entropy $H(C|\mathbf{Y}) = -\sum_c p(C=c)\sum_y p(Y=y|C=c)\log p(C=c|Y=y)$ as a measure of the remaining uncertainty in C once the expression levels Y are known. We then used this measure and a greedy procedure to select multiple disjoint gene sets, $Y_1, \ldots, Y_k$ such that for each set Y_i, $H(C|Y_i)<\eta$ (we set $\eta = 0.5$). In the greedy procedure, we select genes one at a time, and with each selected gene recompute the entropy given the genes already selected in the current set. Once a set is complete (the remaining conditional entropy is below the threshold η), we add all the genes to the selected set, and repeat the procedure (excluding all the selected genes from consideration). We stop when the number of selected genes has reached a user-defined threshold, set by the number of genes feasible for the experimental assay.

To select a time point, we used the same approach. Here, we measured entropy under all time points for multiple randomly selected gene sets of several sizes and plotted the average entropy for each time point (see Amit et al., 2009). We chose the time point with the minimal entropy (i.e., 6 hr post-simulation).

nCounter Data Analysis

After normalization by internal Nanostring controls (spike-normalization following manufacturer's instructions), we normalized the data relying on three control genes (*Ndufa7*, *Tbca*, *Tomm7*) that are the least affected by shRNAs and LPS stimulation. Next, we log-transformed the expression values (Bengtsson and Hossjer, 2006). Five signature genes (*Cxcl5*, *Fos*, *Fst*, *Ereg*, and *Egr2*) that were highly variable across control shRNA samples were removed from subsequent analysis. To score target genes whose expression is significantly affected by perturbations, we used a fold threshold corresponding to an FDR of 2%. For a given shRNA perturbation, a target gene was called as significantly affected when the ratio of the log-expression of this gene upon shRNA knockdown to

the average log-expression of this gene in control shRNA samples was below (or above) a threshold (1/threshold), chosen such that, on average, no more than 2 hits (out of 128 genes in the Nanostring codeset) per control shRNA sample were called. Heatmaps and distance matrix analyses were generated using the software Gene-E (http://www.broadinstitute.org/cancer/software/GENE-E/).

Microarray Hybridization and Processing

For oligonucleotide microarray hybridization, 1 μg of RNA were labeled, fragmented, and hybridized to an Affymetrix Mouse Genome 430A 2.0 Array. After scanning, the expression value for each gene was calculated with RMA (Robust Multi-Array) normalization. The average intensity difference values were normalized across the sample set. Probe sets that were absent in all samples according to Affymetrix flags were removed. All values below 50 were floored to 50.

Detection of Regulated Signaling Genes

To identify differentially regulated signaling components (i.e., kinases, phosphatases, and signaling adaptors or scaffolds), we defined regulated probesets for each condition (TLR agonist) as probesets displaying at least 1.7-fold up- or downregulation in both duplicates of at least one time point, compared to unstimulated controls, using our previously published microarray dataset available in the NCBI Gene Expression Omnibus under the accession number GSE17721 (Amit et al., 2009). Differentially regulated probesets were intersected with lists of kinases, phosphatases, and signaling adaptors and scaffolds. These gene sets were generated combining data from publicly available databases: Panther (http://www.pantherdb.org), Gene Ontology (http://www.geneontology.org), and DAVID (http://david.abcc.ncifcrf.gov). Regulated signaling genes were hierarchically clustered using the Cluster software (Eisen et al., 1998).

Antiviral versus Inflammatory Gene Enrichment

Genes whose expression changed upon BI 2536 treatment in microarrays (Table S4) were evaluated for their enrichment with genes involved in the antiviral and inflammatory programs. When multiple probesets were available for a given gene on the microarray, only the probeset with maximum value was kept for analysis. Thus, the complete microarray consisted of 14,088 genes, among which 804 and 550 genes were identified as part of antiviral and inflammatory programs, respectively (Amit et al., 2009). We performed a hypergeometric test on genes whose expression changed at least 3-fold upon BI 2536 treatment compared to vehicle control (DMSO), in either LPS or poly(I:C) samples. In addition, genes whose expression changed upon BI 2536 treatment in microarrays in response to LPS and/or poly(I:C) stimulation were analyzed for enrichment of Gene Ontology (GO) processes and canonical pathways from curated databases using the Molecular Signature Databse (MSigDB; http://www.broadinstitute.org/gsea/msigdb/index.jsp).

Nanowire-Mediated Drug Delivery and Microscopy

BMDCs were plated on top of etched silicon nanowires (Si NWs) coated with small molecules (Shalek et al., 2010). After 24 hr, cells were stimulated with LPS or poly(I:C), then fixed in 4% formaldehyde in PBS (RT, 10 min). After fixation, each sample was permeabilized with 0.25% Triton X-100 in PBS (RT, 10 min), incubated with Image-iT FX Signal Enhancer (RT, 30 min), and then blocked with 10% goat serum and 0.25% Triton X-100 in PBS (RT, 1 hr). After washing, the samples were placed in 3% IgG-Free BSA & 0.25% Triton X-100 in PBS that contained primary antibodies against either IRF3 or NF-κB P65 (1:175 dilution) and then rocked overnight at 4°C. The following day, the samples were washed with PBS and then incubated with an Alexa Fluor labeled secondary antibody (1:250 dilution) in 3% IgG-Free BSA & 0.25% Triton X-100 in PBS (RT, 60 min). After washing with PBS, the samples were counterstained with 300 ng/ml of DAPI in PBS (RT, 30 min). For each experiment, every stimulus-molecule combination was prepared in triplicate. Once fixed, samples were imaged using an upright confocal microscope (Olympus). For each captured image, the nuclear fraction of the fluorescent protein was calculated after identifying nuclear boundaries using DAPI. Finally, distributions for this quantity across different conditions were compared using a one-way ANOVA analysis.

In Vivo BI 2536 Experiments in a VSV Infection Model

Eight-week-old C57BL/6 male mice (from Charles River Laboratories) received 500 μg of BI 2536 (or vehicle) intravenously, and 50 μg into the footpad 3 hr before and 2 hr after infection with 10^6 pfu of VSV, as previously described (Iannacone et al., 2010), into the footpad. Mice were sacrificed 6 hr post-infection and the draining popliteal lymph nodes were harvested in RNAlater solution (Ambion) before subsequent RNA analysis. All experimental animal procedures were approved by the Institutional Animal Committees of Harvard Medical School and IDI. All infectious work was performed in designated $BL2^+$ workspaces, in accordance with institutional guidelines, and approved by the Harvard Committee on Microbiological Safety.

Microwestern Arrays

The MWA method previously described (Ciaccio et al., 2010) was modified to accommodate a larger number of lysates. The lysates were printed in a ''double-block'' format with each MWA being 18 mm wide by 9 mm long. Twelve samples plus protein marker (Li-Cor 928-40000) were printed with a noncontact piezoelectric arrayer (GeSiM NP2) along the top edge of the block, each block printed forty-eight times on the acrylamide gel. The deck layout is included in Figure S7A. Electrophoresis, transfer, and antibody incubation were performed as previously described with the exception of using a modified 48-well gasket (The Gel Company MMH96) manually cut to have a larger block size in order to isolate antibodies on the nitrocellulose membrane per printed block. The antibodies used in

this study were against β-ACTIN, GAPDH, β-TUBULIN, IκBα (clone L35A5), P65 (clone C22B4), STAT1, p-ABL(C-) (Y245), p-AKT (S473), p-AKT1/2/3 (T308), p-ATF2 (T71), p-ERK1/2 (T202/Y204), p-IKBALPHA (S32), p-IKKA/B (S176/180), p-IRF3 (S396), p-MAPKAPK2 (T222), p-MEK(1/2) (S217/221), p-MET (Y1234/1235), p-P38 (T180/Y182), p-P65 (S536), p-P70S6K (S371), p-P70S6K (T389), p-P90RSK (S380), p-PI3K P85(Y458) P55(Y199), p-PKCD (Y311), p-SAPK/JNK (T183/Y185), p-SEK1/MKK4 (T261), p-STAT1 (S727), p-STAT1 (Y701), p-STAT3 (S727). All antibodies were from Cell Signaling Technology, except for β-ACTIN which was from Santa Cruz Biotechnology. Band intensities were quantified using Li-Cor Odyssey analysis software (V3.0). Circles were applied to the appropriate band on the scanned image and the net intensity was calculated by subtracting the background intensity from the trimmed mean intensity of each band. The net intensity was divided by the average net intensities of β-actin to control for lysate protein concentration. Fold Change was then calculated in relation to time of inhibitor application (time zero).

Phosphotyrosine Peptide Analysis

Tyrosine-phosphorylated peptides were prepared using a PhosphoScan Kit (Cell Signaling Technology) as previously described (Rush et al., 2005). Briefly, 100 million cells were lysed in lysis buffer (20 mM HEPES, 25 mM sodium pyrophosphate, 10 mM beta-glycerophosphate, 9 M urea, 1 mM ortho-vanadate, 1 Roche Ser/Thr phosphatase inhibitor tablet) assisted by sonication on ice using Misonix S-4000 sonicator with five 30 s bursts at 4 W. Lysates were precleared by centrifugation for 15 min at 20,000 g. Approximately 10 mg of total proteins from each SILAC label were mixed, reduced with 10 mM dithiothreitol and alkylated with 25 mM iodoacetamide. After 4-fold dilution 200 μg sequencing grade modified trypsin (Promega, V5113) was added in an enzyme to substrate ratio of 1:100. The total peptide mixtures were then desalted using a tC18 SepPak cartridge (Waters, 500 mg, W AT036790) and resuspended in IAP buffer (50 mM MOPS/NaOH [pH 7.2], 10 mM Na2HPO4, 50 mM NaCl). Peptide immunoprecipitation was performed with protein-G agarose bead-bound anti-phosphotyrosine antibodies pY100. Peptides captured by phosphotyrosine antibodies were eluted under acidic conditions (0.15% trifluoroacetic acid). The IP eluate was analyzed by data-dependent LC-MS/MS using a Thermo LTQ-Orbitrap instrument.

Global Serine, Threonine, and Tyrosine Phosphorylation Analysis

Quantitative analysis of serine-, threonine-, and tyrosine-phosphorylated peptides was performed essentially as described (Villén and Gygi, 2008) with some modifications. After stimulation, cells were lysed for 20 min in ice-cold lysis buffer (8 M Urea, 75 mM NaCl, 50 mM Tris [pH 8.0], 1 mM EDTA, 2 μg/ml Aprotinin (Sigma, A6103), 10 μg/ml Leupeptin (Roche, #11017101001), 1 mM PMSF, 10 mM NaF, 2 mM Na3VO4, 50 ng/ml Calyculin A (Calbiochem, #208851), Phosphatase inhibitor cocktail 1 (1/100, Sigma, P2850) and Phosphatase inhibitor cocktail 2 (1/100, Sigma, P5726). Lysates were precleared by centrifugation at 16,500 g for 10 min and protein concentrations were determined by BCA assay (Pierce). We obtained 3 mg total protein per label out of 30 million cells. Cell lysates were mixed in equal amounts per label and proteins were reduced with 5 mM dithiothreitol and alkylated with 10 mM iodoacetamide. Samples were diluted 1:4 with HPLC water (Baker) and sequencing-grade modified trypsin (Promega, V5113) was added in an enzyme to substrate ratio of 1:150. After 16 hr digest, samples were acidified with 0.5% trifluoroacetic acid (final concentration). Tryptic peptides were desalted on reverse phase tC18 SepPak columns (Waters, 500 mg, WAT036790) and lyophilized to dryness. Peptides were reconstituted in 500 μl strong cation exchange buffer A (7 mM KH_2PO_4, pH 2.65, 30% MeCN) and separated on a Polysulfethyl A column from PolyLC (250 × 9.4 mm, 5 μm particle size, 200 A pore size) using an Akta Purifier 10 system (GE Healthcare). We used an 80 min gradient with a 20 min equilibration phase with buffer A, a linear increase to 30% buffer B (7 mM KH2PO4, pH 2.65, 350 mM KCL, 30% MeCN) within 33 min, 100% B for 7 min and a final equilibration with Buffer A for 20 min. The flow rate was 3 ml/min and the sample was injected after the initial 20 min equilibration phase. Upon injection, 3 ml fractions were collected with a P950 fraction collector throughout the run. 60 fractions were collected of which 3–4 adjacent fractions were combined to obtain 12 samples. Pooling of SCX fractions was guided by the UV214-trace and fractions were combined starting where the first peptide peak appeared. The 12 samples were desalted with reverse phase tC18 SepPak columns (Waters, 100 mg, WAT036820) and lyophilized to dryness. SCX-separated peptides were subjected to IMAC (immobilized metal affinity chromatography) as described (Villén and Gygi, 2008). Briefly, peptides were reconstituted in 200 μl IMAC binding buffer (40% MeCN, 0.1% FA) and incubated for 1 hr with 5 μl of packed Phos-Select beads (Sigma, P9740) in batch mode. After incubation, samples were loaded on C18 StageTips (Rappsilber et al., 2007), washed twice with 50 μl IMAC binding buffer, and washed once with 50 μl 1% formic acid. Phosphorylated peptides were eluted from the Phos-Select resin to the C18 material by loading 3 times 70 μl of 500 mM K_2HPO_4 (pH 7.0). StageTips were washed with 50 μl of 1% formic acid to remove phosphate salts and eluted with 80 μl of 50% MeCN/0.1% formic acid. Samples were dried down by vacuum centrifugation and reconstituted in 8 μl 3% MeCN/0.1% formic acid.

NanoLC-MS/MS Analysis

All peptide samples were separated on an online nanoflow HPLC system (Agilent 1200) and analyzed on a LTQ Orbitrap Velos (Thermo Fisher Scientific) mass spectrometer. Four microliters of peptide sample were autosampled onto a 14 cm reverse phase fused-silica capillary column (New Objective, PicoFrit PF360-75-10-N-5 with 10 μm tip opening and 75 μm inner diameter) packed in-house with 3 μm ReproSil-Pur C18-AQ media (Dr. Maisch GmbH). The HPLC setup was connected via a custom-made electrospray ion source to the mass spectrometer. After sample injection, peptides were separated at an analytical flowrate of 200 nl/min with an 70 min linear gradient (~0.29%B/min) from 10% solvent A (0.1% formic acid in water) to 30% solvent B (0.1% formic

acid/90% acetonitrile). The run time was 130 min for a single sample, including sample loading and column reconditioning. Data-dependent acquisition was performed using the Xcalibur 2.1 software in positive ion mode. The instrument was recalibrated in real-time by coinjection of an internal standard from ambient air (“lock mass option”) (Olsen et al., 2005). Survey spectra were acquired in the orbitrap with a resolution of 60,000 and a mass range from 350 to 1750 m/z. In parallel, up to 16 of the most intense ions per cycle were isolated, fragmented and analyzed in the LTQ part of the instrument. Ions selected for MS/MS were dynamically excluded for 20 s after fragmentation. For the second biological replicate analysis peptides observed to be regulated in the first analysis were loaded into a global parent mass inclusion list and 4 MS/MS scans were reserved for precursors from the inclusion list whereas 12 were performed on the most intense ions per duty cycle.

Identification and Quantification of Peptides and Proteins

Mass spectra were processed using the Spectrum Mill software package (Agilent Technologies) v4.0 b that includes in-house developed features for SILAC-based quantitation and phoshosite localization and also with the MaxQuant software package (version 1.0.13.13) (Cox and Mann, 2008), which was used in combination with a Mascot search engine (version 2.2.0, Matrix Science). For peptide identification in Spectrum Mill an International Protein Index protein sequence database (IPI version 3.60, mouse) was used which was reversed on-the-fly at search time. In MaxQuant a concatenated forward and reversed IPI protein sequence database (version 3.60, mouse) was queried. The mass tolerance for precursor ions and for fragment ions was set to 7 ppm and 0.5 Da, respectively. Cysteine carbamidomethylation was searched as a fixed modification, whereas oxidation on methionine, N-acetylation (Protein) and phosphorylation on serine, threonine, or tyrosine residues were considered as variable modifications. The enzyme specificity was set to trypsin and cleavage N-terminal of proline was allowed. The maximum of missed cleavages was set to 3. For peptide identification the maximum peptide FDR was set to 1%. The minimum identification score was to 5 in Spectrum Mill and to 10 in MaxQuant. SILAC ratios were obtained from the peptide export table in Spectrum Mill and the evidence table in MaxQuant. Arginine to Proline conversion was determined to be 3.42% and 5.55% for both biological replicates, respectively. The conversion was calculated by defining Arg-10 as a fixed modification and by quantifying the ratio between peptides containing normal L-proline (Pro0) and 13C5-15N1-labeled proline (Pro6) with MaxQuant. Each peptide SILAC ratio was corrected for arginine to proline conversion by the formula r[c] = r[o]/((1-p)^n), where r[c] is the corrected ratio, r[o] the observed ratio, p the conversion rate and n the number of proline residues per peptide. The median ratios of all nonphosphorylated peptides were used to normalize the M/L and H/L ratios of all phosphorylated peptides. To allow better peptide grouping, phosphosite localization information obtained from SpectrumMill and MaxQuant were further simplified. Probability scores greater or equal 0.75 were called fully localized and designated with (1.0), scores smaller 0.75 and greater or equal to 0.5 were called ambiguously localized and designated with (0.5), whereas scores smaller than 0.5 were called non-localized and the total number of phosphorylation sites per peptide was designated with an underscore after the peptide sequence. Median SILAC ratios of phosphopeptides for each experiment were calculated over all versions of the same peptide including different charge states and methionine oxidation states. The highest scoring versions of each distinct peptide were reported per experiment. Overlapping data between SpectrumMill and MaxQuant as well as between different biological replicates was analyzed for discrepancies by calculating the mean and standard deviation over all residuals for different ratios of the same phosphopeptide. Residuals were calculated by subtracting the two values for each phosphopeptide derived by SpectrumMill or MaxQuant as well as by two different biological replicates. All peptides were filtered from the data set that had residuals greater than 3 standard deviations distant from the mean as they were not reproducible between two biological replicates or between SpectrumMill and MaxQuant. Data derived from both software packages was combined and MaxQuant data was reported when the same phosphopeptide was identified and quantified by both programs. Log_2 phosphopeptide ratios of BI 2536 treated versus untreated dendritic cells followed a normal distribution that was fitted using least-squares regression. Mean and standard deviation values derived from the Gaussian fit were used to calculate p values. An FDR-based measure was used to assess significance of the findings (Storey and Tibshirani, 2003).

SUPPLEMENTAL REFERENCES

Amit, I., Garber, M., Chevrier, N., Leite, A.P., Donner, Y., Eisenhaure, T., Guttman, M., Grenier, J.K., Li, W., Zuk, O., et al. (2009). Unbiased reconstruction of a mammalian transcriptional network mediating pathogen responses. Science *326*, 257–263.

Bengtsson, H., and Hössjer, O. (2006). Methodological study of affine transformations of gene expression data with proposed robust non-parametric multi-dimensional normalization method. BMC Bioinformatics *7*, 100.

Ciaccio, M.F., Wagner, J.P., Chuu, C.P., Lauffenburger, D.A., and Jones, R.B. (2010). Systems analysis of EGF receptor signaling dynamics with microwestern arrays. Nat. Methods *7*, 148–155.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. *26*, 1367–1372.

Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. Proc. Natl. Acad. Sci. USA *95*, 14863–14868.

Geiss, G.K., Bumgarner, R.E., Birditt, B., Dahl, T., Dowidar, N., Dunaway, D.L., Fell, H.P., Ferree, S., George, R.D., Grogan, T., et al. (2008). Direct multiplexed measurement of gene expression with color-coded probe pairs. Nat. Biotechnol. *26*, 317–325.

Guris, D.L., Fantes, J., Tara, D., Druker, B.J., and Imamoto, A. (2001). Mice lacking the homologue of the human 22q11.2 gene CRKL phenocopy neurocristopathies of DiGeorge syndrome. Nat. Genet. *27*, 293–298.

Hemmeryckx, B., Reichert, A., Watanabe, M., Kaartinen, V., de Jong, R., Pattengale, P.K., Groffen, J., and Heisterkamp, N. (2002). BCR/ABL P190 transgenic mice develop leukemia in the absence of Crkl. Oncogene *21*, 3225–3231.

Hole, K., Clavijo, A., and Pineda, L.A. (2006). Detection and serotype-specific differentiation of vesicular stomatitis virus using a multiplex, real-time, reverse transcription-polymerase chain reaction assay. J. Vet. Diagn. Invest. *18*, 139–146.

Iannacone, M., Moseman, E.A., Tonti, E., Bosurgi, L., Junt, T., Henrickson, S.E., Whelan, S.P., Guidotti, L.G., and von Andrian, U.H. (2010). Subcapsular sinus macrophages prevent CNS invasion on peripheral infection with a neurotropic virus. Nature *465*, 1079–1083.

Inglis, K.J., Chereau, D., Brigham, E.F., Chiou, S.S., Schöbel, S., Frigon, N.L., Yu, M., Caccavello, R.J., Nelson, S., Motter, R., et al. (2009). Polo-like kinase 2 (PLK2) phosphorylates alpha-synuclein at serine 129 in central nervous system. J. Biol. Chem. *284*, 2598–2602.

Lansing, T.J., McConnell, R.T., Duckett, D.R., Spehar, G.M., Knick, V.B., Hassler, D.F., Noro, N., Furuta, M., Emmitte, K.A., Gilmer, T.M., et al. (2007). In vitro biological activity of a novel small-molecule inhibitor of polo-like kinase 1. Mol. Cancer Ther. *6*, 450–459.

Moffat, J., Grueneberg, D.A., Yang, X., Kim, S.Y., Kloepfer, A.M., Hinkle, G., Piqani, B., Eisenhaure, T.M., Luo, B., Grenier, J.K., et al. (2006). A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. Cell *124*, 1283–1298.

Müller, U., Steinhoff, U., Reis, L.F., Hemmi, S., Pavlovic, J., Zinkernagel, R.M., and Aguet, M. (1994). Functional role of type I and type II interferons in antiviral defense. Science *264*, 1918–1921.

Olsen, J.V., de Godoy, L.M., Li, G., Macek, B., Mortensen, P., Pesch, R., Makarov, A., Lange, O., Horning, S., and Mann, M. (2005). Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. Mol. Cell. Proteomics *4*, 2010–2021.

Ong, S.E., and Mann, M. (2006). A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). Nat. Protoc. *1*, 2650–2660.

Rappsilber, J., Mann, M., and Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat. Protoc. *2*, 1896–1906.

Reindl, W., Yuan, J., Krämer, A., Strebhardt, K., and Berg, T. (2009). A pan-specific inhibitor of the polo-box domains of polo-like kinases arrests cancer cells in mitosis. ChemBioChem *10*, 1145–1148.

Rush, J., Moritz, A., Lee, K.A., Guo, A., Goss, V.L., Spek, E.J., Zhang, H., Zha, X.M., Polakiewicz, R.D., and Comb, M.J. (2005). Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. Nat. Biotechnol. *23*, 94–101.

Shalek, A.K., Robinson, J.T., Karp, E.S., Lee, J.S., Ahn, D.R., Yoon, M.H., Sutton, A., Jorgolli, M., Gertner, R.S., Gujral, T.S., et al. (2010). Vertical silicon nanowires as a universal platform for delivering biomolecules into living cells. Proc. Natl. Acad. Sci. USA *107*, 1870–1875.

Shapira, S.D., Gat-Viks, I., Shum, B.O., Dricot, A., de Grace, M.M., Wu, L., Gupta, P.B., Hao, T., Silver, S.J., Root, D.E., et al. (2009). A physical and regulatory map of host-influenza interactions reveals pathways in H1N1 infection. Cell *139*, 1255–1267.

Steegmaier, M., Hoffmann, M., Baum, A., Lénárt, P., Petronczki, M., Krssák, M., Gürtler, U., Garin-Chesa, P., Lieb, S., Quant, J., et al. (2007). BI 2536, a potent and selective inhibitor of polo-like kinase 1, inhibits tumor growth in vivo. Curr. Biol. *17*, 316–322.

Storey, J.D., and Tibshirani, R. (2003). Statistical significance for genomewide studies. Proc. Natl. Acad. Sci. USA *100*, 9440–9445.

Tager, A.M., Kradin, R.L., LaCamera, P., Bercury, S.D., Campanella, G.S., Leary, C.P., Polosukhin, V., Zhao, L.H., Sakamoto, H., Blackwell, T.S., and Luster, A.D. (2004). Inhibition of pulmonary fibrosis by the chemokine IP-10/CXCL10. Am. J. Respir. Cell Mol. Biol. *31*, 395–404.

Villén, J., and Gygi, S.P. (2008). The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry. Nat. Protoc. *3*, 1630–1638.

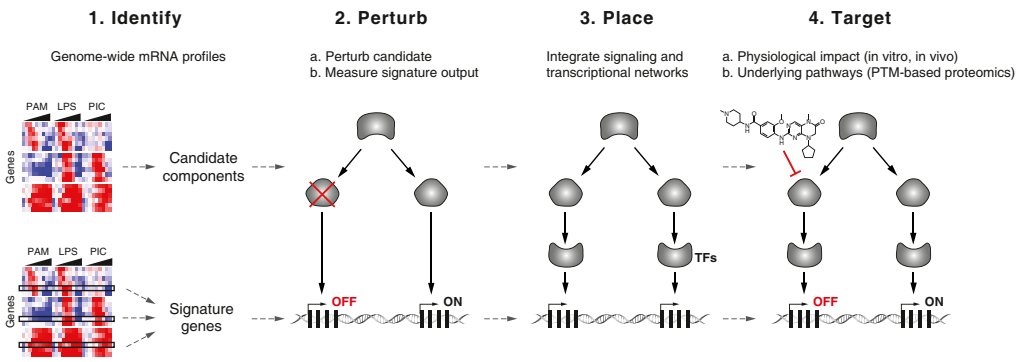

Figure S1. A Systematic Approach to Dissect Signaling Pathways, Related to Figure 1

Shown is a schematic depicting the strategy consisting of four steps (from left to right): (1) extract both candidate signaling regulators and signature genes; (2) perturb each candidate with shRNAs and measure the effect on the expression of signature genes; (3) compare perturbation profiles of signaling and transcriptional regulators to start assembling pathways; (4) use small-molecule targeting of signaling nodes of interest to (a) evaluate the physiological relevance of new signaling node and (b) identify underlying pathways by discovering downstream effector molecules.

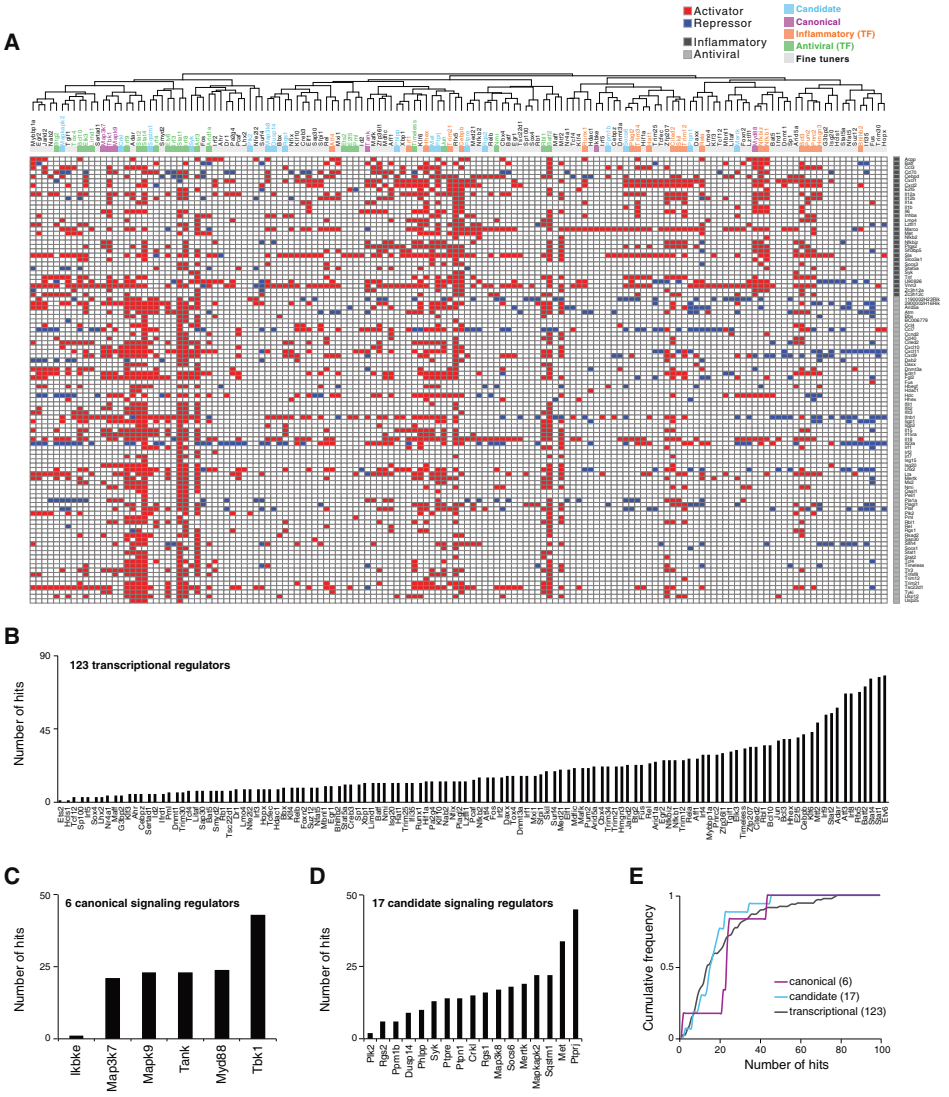

Figure S2. Perturbations of Signaling and Transcriptional Regulators Have Similar Effects on the TLR Signature Genes, Related to Figure 2

(A) Perturbation profiles of 6 canonical (purple) and 17 candidate (light blue) signaling regulators and 123 transcriptional regulators (TF) partitioned into regulators of the inflammatory (orange) and antiviral (green) programs, and fine tuners (gray), as previously defined in Amit et al. (2009). Shown are the perturbed regulators (columns) and their statistically significant effects (FDR < 2%) on each of the 118 TLR signature genes (rows). Red: significant activating relation (target gene expression decreased following perturbation); blue: significant repressing relation (target gene expression increased following perturbation); white: no significant effect. The column on the right indicates whether signature genes belong to the antiviral (light gray) or the inflammatory (dark gray) programs.

(B–D) Shown are the numbers of signature genes hits (y axis, "hits") significantly affected by knockdown of each regulator (x axis) for the regulator categories shown in A: 123 transcriptional (B) and 6 previously known (C) and 17 candidate (D) signaling regulators.

(E) Candidate signaling regulators affect a similar number of "signature" genes compared to transcriptional regulators. Shown is the cumulative distribution of the number of hits for the regulators shown in (B)–(D).

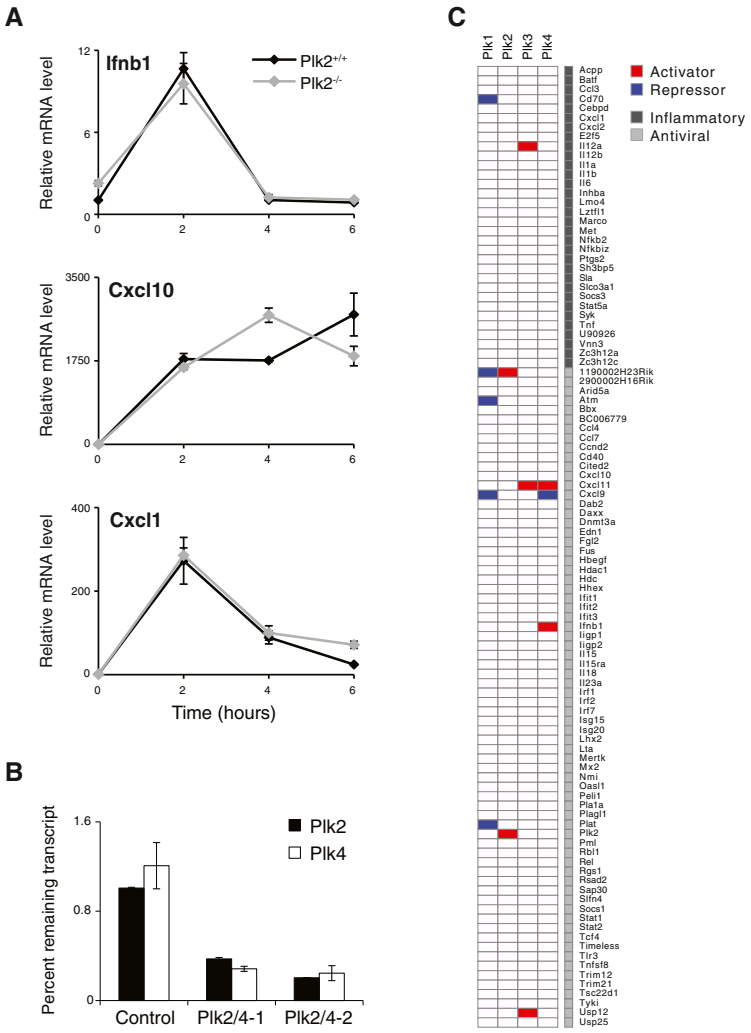

Figure S3. Individual Perturbation of Plk Family Members Does Not Affect TLR Output Gene Expression in DCs, Related to Figure 4

(A) *Plk2*-deficient BMDCs respond to LPS similarly to wild-type cells. Shown are mRNA levels (qPCR; relative to t = 0) for *Ifnb1* (left), *Cxcl10* (middle), and *Cxcl1* (right) in three replicates per time point. Error bars represent the SEM.

(B) Combinatorial knockdown levels of *Plk2* and *4* in BMDCs. Shown are mRNA levels (qPCR), relative to nontargeting shRNAs (Control), of *Plk2* and *4* in BMDCs using two independent combinations of shRNAs (*Plk2/4-1* and *-2*). Three replicates in each experiment; error bars represent the SEM.

(C) Perturbations of individual Plk family members do not affect TLR signature genes. Shown are the perturbed Plks (columns) and their statistically significant effects (FDR < 2%) on each 118 TLR signature genes (rows). Red: significant activating relation (target gene expression decreased following perturbation); blue: significant repressing relation (target gene expression increased following perturbation); white: no significant effect. The column on the right indicates whether signature genes belong to the antiviral (light gray) or the inflammatory (dark gray) programs.

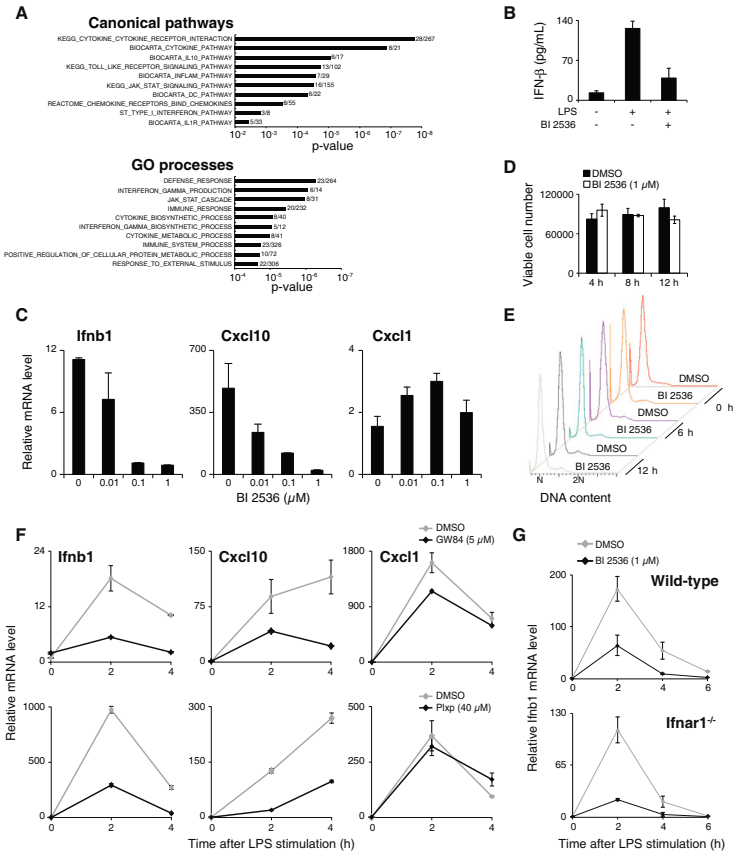

Figure S4. BI 2536-Mediated Plk Inhibition Abrogates Antiviral Cytokine Production at the Protein and mRNA Levels, without Affecting the Viability and Cell-Cycle Status of DCs, Related to Figure 4

(A) Gene enrichment analysis of BI 2536-dependent genes from microarray measurements. Overlaps between the 311 unique genes downregulated 3-fold by BI 2536 treatment upon LPS or poly(I:C) stimulation (Table S4), and Gene Ontology (GO) processes and canonical pathways (including the KEGG, REACTOME, and BIOCARTE databases present in the Molecular Signatures Database [MSigDB]; see Experimental Procedures). Shown are p values (x axis) derived from the overlaps (n/N; top of each bar) between the number of queried genes (n) and genes present in indicated genesets (N).

(B) BI 2536 strongly inhibits IFN-β secretion by BMDCs. Shown is IFN-β protein concentration (y axis; measured by ELISA) in the supernatant of BMDCs treated with DMSO vehicle (–) or BI 2536 (1 μM; +), and stimulated with LPS (+) or left unstimulated (–) for 6 hr. Three replicates in each experiment; error bars are the SEM.

(C) BI 2536 inhibits antiviral cytokine mRNA production in a dose-dependent manner. Shown are mRNA levels (y axis, qPCR; relative to vehicle control treatment) for two antiviral cytokines (*Ifnb1*, *Cxcl10*) and one inflammatory cytokine (Cxcl1) following LPS stimulation in BMDCs pretreated with increasing amounts of BI 2536 (x axis). Three replicates in each experiment; error bars are the SEM.

(D) BMDC viability is unaffected by Plk inhibition with BI 2536. Shown are viable cell numbers (y axis, measured by Alamar blue; relative to a standard curve generated using a range of cell densities) after treatment with BI 2536 (white bars) or DMSO vehicle (black bars) at different time points following addition of BI 2536 (x axis). Three replicates in each experiment; error bars are the SEM.

(E) The cell-cycle state of BMDCs remains unchanged upon Plk inhibition with BI 2536. Shown are DNA contents (flow cytometry) of BMDCs stained with propidium iodide (PI) after treatment with BI 2536 or DMSO vehicle control for 0, 6, and 12 hr.

(F) Plk inhibitors structurally unrelated to BI 2536 also abrogate transcription of mRNAs for antiviral cytokines following stimulation with LPS. Shown are mRNA levels (qPCR; relative to t = 0) for *Ifnb1*, *Cxcl10*, and *Cxcl1* in BMDCs stimulated with LPS and treated with GW843682X (GW84; top) or Poloxipan (Plxp; bottom) (black line), or with DMSO vehicle (gray line) for 1 hr prior to stimulation. Three replicates for each experiment; error bars are the SEM.

(G) Plks are directly downstream of TLR engagement. Shown are Ifnb1 mRNA levels (y axis, qPCR; relative to t = 0) following LPS stimulation for indicated times (x axis) in wild-type (top) and *Ifnar1*$^{-/-}$ (bottom) BMDCs treated with BI 2536 (1 μM; black) or vehicle control (DMSO; gray). Error bars represent the SEM.

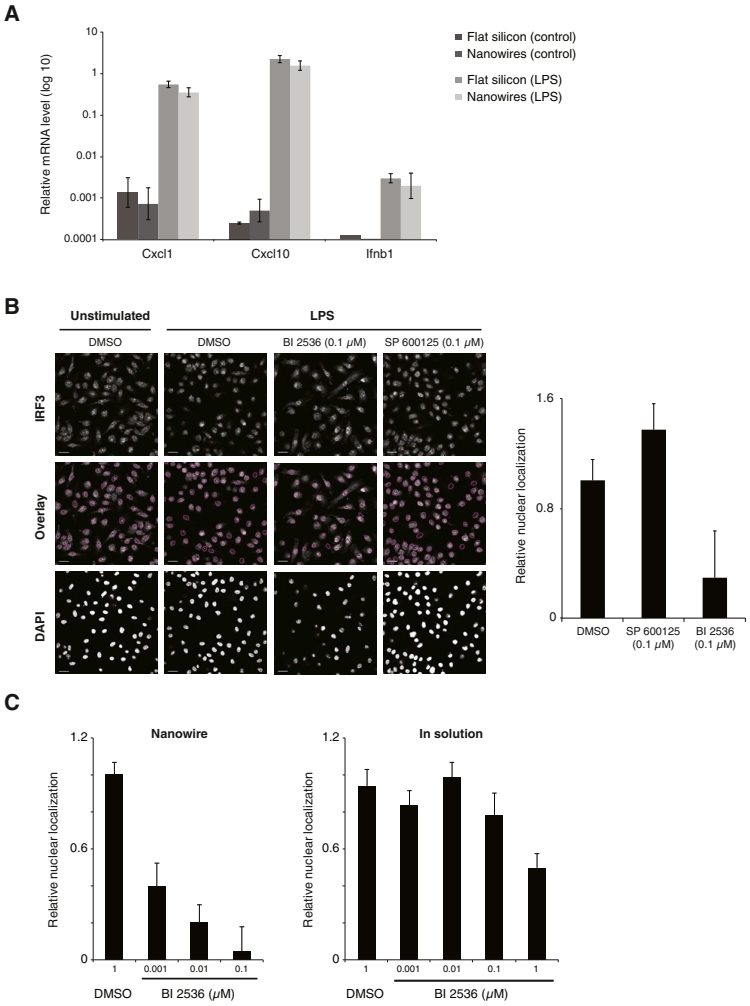

Figure S5. BI 2536-Mediated Plk Inhibition Blocks IRF3 Nuclear Translocation in LPS-Stimulated DCs, Related to Figure 5

(A) DCs plated on vertical silicon NW respond normally to TLR stimulation. Shown are cytokine mRNA levels (qPCR; relative to Gapdh mRNA) in BMDCs plated on NW or a flat silicon surface, and stimulated (LPS) or left untreated (control). Left to right: *Cxcl1*, *Cxcl10*, *Ifnb1*. Three replicates in each experiment; error bars are the SEM.

(B) BI 2536 inhibits IRF3 nuclear translocation following LPS stimulation. Shown are confocal micrographs (left panel) of BMDCs plated on vertical silicon NW precoated with vehicle control (DMSO), Plk inhibitor (BI 2536), or control Jnk inhibitor (SP 600125) and stimulated with LPS for 45 min (reflecting peak time of nuclear translocation for IRF3 in the context of LPS stimulation) or left unstimulated. Cells were analyzed for DAPI and IRF3 staining. Scale bars, 5 μM. Nuclear translocation (from confocal micrographs) of IRF3 was quantified (right panel) using DAPI staining as a nuclear mask (purple circles on micrographs) to determine the ratio of total versus nuclear fluorescence (y axis) in BMDCs cultured on NW coated with BI 2536, SP 600125, or vehicle control (DMSO; x axis). Three replicates in each experiment; error bars are the SEM.

(C) Decrease in IRF3 nuclear translocation may be more efficient with NW-mediated delivery of BI 2536 than with delivery in solution. Shown are quantifications of confocal micrographs from BMDCs plated on vertical NW precoated with different amounts of BI 2536 (Nanowire; left panel) or left blank to allow in-solution delivery of BI 2536 (In solution; right panel). Cells were stimulated with poly(I:C) for 2 hr prior to staining for DAPI and IRF3 as in (B). Error bars represent the SEM.

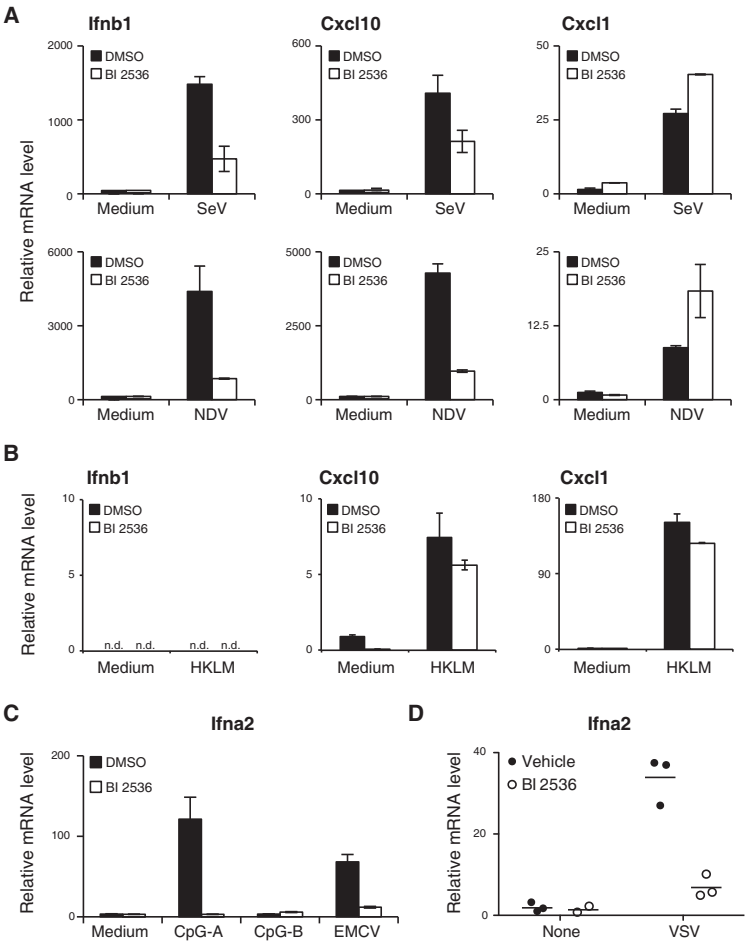

Figure S6. Plks Are Critical in Antiviral Responses In Vitro and In Vivo, Related to Figure 6

(A) Plks are critical in RIG-I-mediated antiviral responses in vitro in DCs. Shown are mRNA levels (qPCR; relative to control, "medium") in conventional DCs (GM-CSF-induced BMDCs) treated with BI 2536 (white bars) or DMSO vehicle (black bars) and infected at a multiplicity of infection (moi) 1 with Sendai virus (SeV; top) or Newcastle disease virus (NDV; bottom). Three replicates in each experiments; error bars are the SEM.

(B) Plk inhibition does not affect DC response to *Listeria monocytogenes*, a natural TLR2 agonist. Shown are mRNA levels (qPCR; relative to t = 0) for *Ifnb1*, *Cxcl10*, *and Cxcl1* in BMDCs stimulated with heat-killed *Listeria monocytogenes* (HKLM; moi 5) and treated with BI 2536 (white bars) or with DMSO vehicle (black bars) for 1 hr prior to stimulation. Three replicates for each experiment; error bars are the SEM.

(C) Plks are critical in type I interferon $\alpha 2$ (*Ifna2*) gene production by pDCs. Shown is the mRNA level (qPCR; relative to control, "medium") of *Ifna2* in pDCs (Flt3L-induced BMDCs) treated with BI 2536 (1 μM; white bars) or DMSO control (black bars) and stimulated with CpG-A or -B or infected with EMCV (moi 10). Three replicates in each experiment; error bars are the SEM.

(D) Plk inhibition in vivo inhibits type I IFN $\alpha 2$ production in the lymph node. Shown is *Ifna2* mRNA level (qPCR; relative to uninfected animals) from popliteal lymph nodes of mice injected with BI 2536 (white circles) or DMSO vehicle (black circles) prior to and during the course of infection with VSV intra-footpad. Nodes were harvested 6 hr post-infection. Each circle represents one animal (n = 3). Data are representative of two or three independent experiments for each condition.

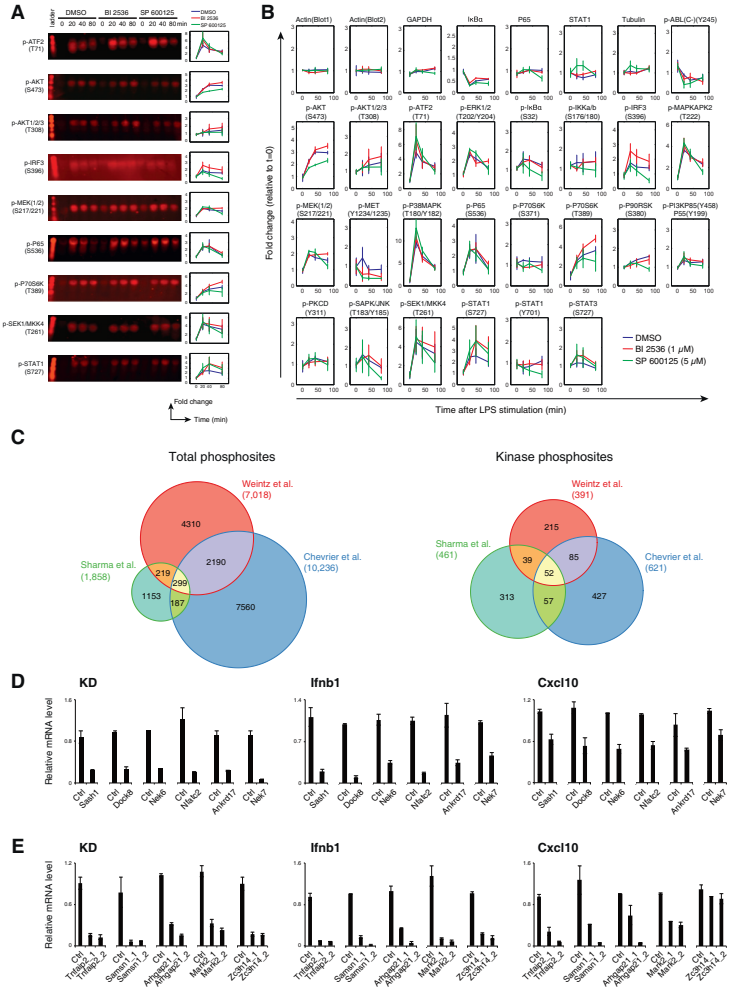

Figure S7. Plk Inhibition Does Not Affect Known TLR Signaling Components but Affects Eleven Newly Identified Plk-Dependent Phosphoproteins, Related to Figure 7

(A and B) BI 2536-mediated Plk inhibition does not affect protein and/or phosphorylation levels of known TLR signaling nodes. (A) Shown are representative MWA (see Experimental Procedures) blots obtained from analyzing lysates from BMDCs pretreated with DMSO, BI 2536 (1 μM), or SP 600125 (5 μM) and stimulated with LPS for 0, 20, 40, 80 min. Blots were analyzed using indicated antibodies (leftmost), and fold change in fluorescence signals was quantified relative to t = 0 (right; see Experimental Procedures). Error bars are the SEM of triplicate MWA blots. (B) Shown are the differential protein and phosphorylation levels (fold change; y axis) of 6 proteins and 23 phosphosites in BMDCs treated with BI 2536 (red line), control JNK inhibitor (SP 600125; green line), or DMSO vehicle (blue line), and stimulated with LPS (0, 20, 40, 80 min; x axis). Band intensities on MWA blots were quantified using Li-Cor Odyssey analysis software (Experimental Procedures). For each antibody, data were normalized to β-actin levels; error bars are the SEM of triplicate MWA blots.

(C and D) Eleven Plk-dependent phosphoproteins are critical for TLR-mediated antiviral responses in DCs. Shown are mRNA levels (qPCR; relative to non-targeting control shRNAs, Ctrl) for knockdown (KD) efficiency (left), *Ifnb1* (middle), and *Cxcl10* (right) in BMDCs following LPS stimulation. Genes with one and two shRNAs are shown in (C) and (D), respectively. Three replicates in each experiment; error bars are the SEM.

(E) Comparison of phosphosites identified in our study and in two recent reports (Weintz et al., 2010 and Sharma et al.. 2010). Shown are proportional Venn diagrams of the total unique phosphosites identified by the 3 studies (left), and the phosphosites harbored by kinases only (right). Total numbers of unique phosphosites per study are indicated in parentheses. Error bars are the SEM.